High latitude thermosphere meso-scale studies and long-term database investigations with the new Scanning Doppler Imager and Fabry-Perot Interferometers

Ho Ching Iris Yiu

Thesis presented for the degree of Doctor of Philosophy

Atmospheric Physics Laboratory

Department of Physics and Astronomy

University College London UCL

May 2014

I, *Ho Ching Iris Yiu*,

confirm that the work presented in this thesis is my own.

Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

In loving memory of my father, Yiu Cheuk Wah

Abstract

The SCANning Doppler Imager, SCANDI, is a new all-sky development on the network of UCL Fabry Perot Interferometers, FPI, in the Arctic region. It is able to provide high temporal and spatial resolution measurements on the upper atmosphere, in scales of minutes and tens of kilometres. This is a step change from the single look direction provided by a standard FPI. SCANDI's mechanics, control-software, data analysis and calibration methods are presented. This includes details on techniques that are specific to the instrument's all-sky capability.

The meso-scale data are of particular interest to developments in atmospheric models, which look for increasingly small-scale structures. An ionospheric cusp study utilising SCANDI is able to provide direct neutral measurements at high spatial resolution. It indicates Joule heating is likely to be one of the main sources which contributes to atmospheric cusp density upwelling. The neutral cusp results also complement the CHAMP satellite density data and are able to demonstrate the neutral thermosphere has a considerable meso-scale structure reflecting its response to the cusp dynamics.

The temperature variation over the recent solar-cycle in years 1999-2009 is analysed, which covers the extended solar minimum of cycle 23/24. It is also the longest continual period of FPI observation from a single site that has been analysed. The dataset demonstrates that the polar cap neutral temperature is unaffected by the unusually low solar activity, possibly because of high latitude geomagnetic contributions. The difference between measurements and MSIS model results are explored, which demonstrate MSIS unexpectedly over-estimated the neutral temperature by 200 K at this solar maximum. It indicates MSIS is unable to recreate the low temperatures in this unusually weak solar cycle.

A new technique was also developed for the study of FPI long term neutral temperature by modelling the red and green-line emission profiles. It is demonstrated that the FPI-measured temperature can be lowered by $\sim$50K because of variations in the width of the emission height profile and peak altitude. This needs to be considered in long-term temperature trend analysis, especially where looking for trends of as small as a few Kelvin.

Acknowledgements

I wish to first acknowledge and thank the APL group, particularly the FPI instrument team. This includes my supervisor, Dr. Anasuya Aruliah, for providing valuable guidance throughout my PhD. As well as being patient and understanding especially during the thesis writing process. This thesis would not be possible without all the help she has given. Dr. Ian McWhirter, who always offers expert answers to the sometimes silly questions, for providing thesis advice in his own free time and making the Arctic instrumentation campaigns some of the most enjoyable trips in my life. Dr. Eoghan Griffin, for helping me to become part of the FPI team when I first started and for the useful advice he has given throughout my time at UCL. Dr. Andy Charalambous, who is always friendly and positive and encourages me to be involved in the mechanical side of the instrumentation. Furthermore, my secondary supervisor, Prof. Alan Aylward, has offered useful advice on my research. My thesis office mates, Jo, Will and Dayle, have also helped me through the many thesis-related problems and panic.

I would like to further acknowledge the solar terrestrial scientists that I have met and worked with throughout campaigns, conferences and collaborations. Particularly Prof. Mike Kosch, for our collaborations in the EFOLD and tri-static campaigns and Dr. Björn Gustavsson, for providing important advice on modelling airglow and aurora. I also wish to thank the RAL EISCAT team for patiently providing answers to my data analysis questions. Particularly Dr. Vikki Howells, who offered valuable help in the EISCAT campaigns. Prof. Herb Carlson who provided both extra data and expert advice on the cusp study and responded to emails when he was extremely busy himself.

This thesis would not be possible without my family, especially my parents who have supported me throughout my valuable education. I am thankful for the care and advice my parents and my grandparents have provided throughout my life. Alun, who builds up my confidence at difficult times and gives me the much-needed motivation and always makes me relax and smile even with his hectic schedule of exams and work. My friends, Crystal, Vicky, Kate, Dee and Jenny, have provided very much appreciated distractions when appropriate, but always understand when I am absent in the past year. Finally, Mrs. Isabel Carmichael, my secondary school Physics teacher, who inspired and encouraged me to study Physics at university. I would not be involved in Physics research without her.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Structure and Overview

The new SCAnning Doppler Imager (SCANDI) is a major topic of this thesis, which is a recent addition to the original network of University College London (UCL) single-viewing Fabry Perot Interferometers (FPIs). An overview of the instrumentation is provided, the details of data analysis and calibration processes are also discussed. A number of topics are shown which apply data from SCANDI and the UCL FPI database, which includes a solar-cycle temperature trend study that covers the recent extended solar minimum, cycle 23/ 24. FPI data are used to investigate the long-term temperature trend. Finally, a study of average neutral mechanisms involving the cusp is presented, which is supported by three individual cusp case studies. The high spatial resolution from SCANDI is particularly useful for examining small-scale thermospheric structures that develop in the cusp region.

The FPI temperature trend study provides the longest neutral temperature analysis of the F-region polar cap region. The additional analysis of the effect of changing emission height on the FPI-measured temperature trend is also new, whereas previous polar cap FPI studies usually assume a fixed FPI measuring altitude, or are individual case studies. The extended solar minimum is also an unique and important dataset. The FPI temperatures are further compared to two models, MSIS and CMAT2, in great detail. Direct neutral measurements at high spatial and time resolutions are sparse and the FPI data here are valuable in future neutral atmospheric model developments. The measurements are able to provide a more accurately modelled neutral atmosphere at higher resolution, for example, using high spatial resolution embedded grids. The cusp study is the first cusp analysis that utilises direct neutral measurements from the FPI and SCANDI. Previous

experiments usually concentrate on the magnetospheric and ionospheric effects, where neutral contributions are assumed to be minimal. The average effects of the neutral atmosphere are provided for solar maximum and minimum conditions, which is supported by three case studies representing the different geomagnetic conditions.

1.2 Aim of the thesis and scientific objectives

A new instrument, the SCAnning Doppler Imager (SCANDI), developed by the Atmospheric Physics Laboratory, University College London (APL-UCL), is presented first. It was deployed at Longyearbyen, Svalbard (78^0N, 16^0E) in December, 2006, and is an all-sky version of the already existing UCL Fabry-Perot Interferometer (FPI). A network of FPIs has been observing the high latitude northern hemispheric regions of upper atmosphere since 1981.

Detailed descriptions of the Fabry-Perot etalon and its physics/optics have been documented well in e.g. McWhirter [1993] and others, hence only an overview is provided in this thesis. The Fabry-Perot etalon is a very high resolution interferometer and can detect very small changes in the wavelength of light. The high resolution FPI is ideal for the detection of tiny Doppler shifts observed in the thermospheric emission lines of interest. Furthermore, the new SCANDI is able to provide higher spatial and temporal resolution measurements, which are well suited for the detection of meso-scale behaviour. An overview of SCANDI instrumentation is provided, and data analysis and calibration processes are also described.

A ten-year temperature trend study is then presented, taken from the 25-year Svalbard FPI database. Though previous FPI neutral temperature studies in high latitude regions are available e.g. Killeen et al. [1995], they are shorter data sets involving approximately half a solar cycle, and/or the study concentrates on wind instead of temperature data (e.g. Aruliah et al. [1996]). It is therefore the longest directly-measured F-region neutral temperature dataset that has been presented as of 2011. Furthermore, the ten years of FPI data here cover the recent extended solar minimum (cycle 23/24) which is unique, with reported low mass density [Emmert et al., 2010] between altitudes of 200-600km and temperatures [Heelis et al., 2009]. The average FPI temperature results are also compared against model results, and the data correlation with Kp/Ap and F10.7 indices are investigated, following work from Killeen et al. [1995]. The relationship between neutral temperature and additional indices, auroral electrojet (AE) and an alternative

EUV solar proxy S10.7 [Dudok de Wit and Bruinsma, 2011, Tobiska et al., 2008, Floyd et al., 2002], are also tested. Together with the emission profile modelling work, the 10 years of data can provide insight into the application of FPI neutral temperature data on long term thermospheric cooling, possibly relating to the increase in greenhouse gases.

As part of the temperature trend chapter, the red (630.0nm) and green (557.7nm) line atomic oxygen emissions are also studied in this thesis. They are the two main emissions observed by the FPI. Most FPI analysis assumes fixed peak emission altitudes, and the neutral parameters measured are taken as representative of the neutral atmosphere at fixed heights. The fixed altitudes are typically ∼240km and ∼110km for red and green-line respectively, with about 20km adjustments according to the average solar activity level. The actual height of the peak emission can then be estimated by matching FPI-measured neutral temperature with modelled temperature at a certain height. This height is then assumed to be the peak emission altitude. The actual FPI observations are height-integrated measurements of parameters that are weighted by the emission profiles. Three simple airglow models are tested and compared, utilising EISCAT ionospheric measured and MSIS neutral atmospheric modelled input. The average results of the models, which are generally similar, are then applied to estimate the effect of emission profile changes on FPI observations. One of the limitations of the airglow model in this thesis is that the model is not designed for auroral particle precipitation periods. However, with consideration that the atmosphere is only under high activity conditions of Kp greater than 3+ for about 10% of the time e.g. Foster et al. [1981, 1986], the airglow model should still be able to provide an average representation of the emission profile, and therefore its effect on the FPI-measured parameters. This is especially important for long term neutral temperature trend studies that look for small temperature variations of the order of 5K/year.

Finally, the averaged FPI temperature data are then applied to a cusp investigation. The thermospheric cusp is known to pass over Svalbard and there is renewed interest in neutral mechanisms involving the thermospheric contribution to cusp upwelling mechanisms. This is because of recent studies based on satellite measurements, which detected neutral upwelling at high altitudes (300km or above) at the cusp [e.g. Clemmons et al. [2008], Lühr et al. [2004]. There are debates on the exact mechanism that is required to sustain the large neutral density bulges. E-region joule heating and direct heating from soft electron precipitation are some of the suggestions. Furthermore, it has been suggested that the neutral atmosphere is likely to play a role in cusp density and patch formation

[Carlson, 2007]. The addition of measured neutral dynamics to the ionospheric and magnetospheric detections provides new information on the role of the neutral atmosphere under cusp conditions. The ten years of FPI data along the meridional direction are used to contrast the average thermospheric cusp footprint over the recent solar cycle, with a comparison of solar maximum and minimum results. Three case studies under different Interplanetary Magnetic Field (IMF) and geomagnetic conditions are also presented in support of the average data. Furthermore, high spatial resolution data from SCANDI from the last three years of the data set are investigated in detail. They give the position of the optical cusp relative to regions of elevated temperatures, as well as regions of upwelling from the zenith wind measurements. The dataset therefore provides new information on the cusp upwelling mechanism from the perspective of the neutral atmosphere's influence.

1.3 A statement of original research

The work that is included in this thesis was undertaken by the author, unless otherwise stated. The SCANDI instrumentation chapter is predominantly focused on data analysis and calibration studies led by the author. It includes an overview of instrumentation and software development and implementation led by Dr. Ian McWhirter. The first SCANDI paper, describing instrumentation and analysis was published by Aruliah et al. [1996]. The science chapters involving emission model, temperature trend and cusp studies are also all original work led by the author of this thesis.

Chapter 2

Scientific Background

The high latitude polar regions on Earth are investigated in this thesis for altitudes above ∼100km. This region is an area of the atmosphere called the thermosphere. At high latitudes, the neutral thermosphere is highly coupled to the ionosphere, which is the charged component of the atmosphere formed mostly by solar (extreme) ultra-violet ionisations, at wavelengths of 121 to 10 nm and photons of energy 10.2 to 124 eV. The ionosphere is then connected by geomagnetic field lines to the magnetic outer regions of the atmosphere, called the magnetosphere. The magnetosphere can be further divided to an inner region which rotates daily with the Earth, and an outer region that circulates under direct influence from the solar wind.

The UCL FPIs provide vital observations of an altitude region of the neutral thermosphere that is not normally provided by other instrumentation. The altitudes they observe (∼110km and ∼240km) are too high for direct balloon measurements and too low for a stable satellite orbit [McWhirter, 1993]. The Mass Spectrometer and Incoherent Scatter (MSIS) model [Hedin, 1987, 1991] is a model that is widely used to determine the state of the neutral atmosphere. It is an empirical model based on years of observations from both ground and space (section 2.6.1). Aksnes et al. [2007] points out that there is a lack of neutral temperature measurements of the middle thermosphere between 130 and 250km, because the MSIS thermospheric temperature profile is constrained mostly by density and temperature measurements outside of this range. Hence, the UCL database of 30 years of neutral temperature measurements are valuable. The high spatial resolution data from SCANDI is particularly important because it can give further insight into interesting regions such as the cusp (section 2.3.6).

Furthermore, Aruliah et al. [2010] noted that, owing to the high viscosity of the upper

thermosphere, which has a density of the order 10^{14}-10^{15} particles m^{-3}, the dynamics are predicted to have a large spatial scale at the order of ∼1000 km and long temporal scale of 1-3 hours. However, meso-scale variations have been observed by the FPIs over the years e.g. Aruliah et al. [2004, 2005], and recent studies with SCANDI [Aruliah et al., 2009, 2010, Griffin et al., 2008, 2009] have shown that small-scale neutral dynamics cannot be ignored when describing ion-neutral interactions e.g. within an auroral arc. Small-scale effects are also important for high latitude heating factors such as Joule heating, which is dependent on the difference in ion velocity and neutral wind that can vary at scales of ∼10km and ∼minutes. The small scale variation in the neutral component of the thermosphere is mostly driven by changes in ion velocity. Small-scale and localised heating can also be found, relating to ionospheric/magnetospheric features such as the cusp. This can in turn affect the ionospheric composition, energy and momentum transfer, which then alters the redistribution of magnetospheric energy between Joule heating and acceleration of the neutral gas [Aruliah et al., 1999].

This section describes basic science that is relevant for the high latitude thermosphere, with an introduction to the magnetosphere. The magnetospheric section focuses on the magnetosphere's influence on the ionosphere and neutral thermosphere. The dynamics and chemistries for both the charged and neutral components of the thermosphere are outlined.

2.1 Basic physics

A few essential physics laws on the property of gas and magnetoplasma are described here, and further details can be found in books and papers, such as Hargreaves [1992] and Schunk and Nagy [2000a].

2.1.1 Thermal equilibrium

The mean neutral temperature is one of the parameters that is measured by the FPIs. The FPIs measure the atomic oxygen emissions at wavelengths of 630.0nm and 557.7nm, and it is assumed that the atomic oxygen has achieved thermo-equilibrium with its surrounding before the emission occurs.

The lifetime between the formation of an excited atomic oxygen in the thermosphere and the emission of photons at 630.0nm wavelength is about 110s and for 557.7nm is

about 1s. The red and green-line emissions are found at the F and E-regions (around 240km and 110km) respectively. As an example, the average ion-neutral collision frequencies are about $1s^{-1}$ in the F-region and $500s^{-1}$ in the E-region. Hence, the lifetimes of both emissions allow the atomic oxygen particles to reach thermal equilibrium with their surroundings, and the temperatures deduced from the FPI measurements of the emissions are representative of the neutral atmospheric temperatures.

In thermal equilibrium, the distribution of velocities of molecules within a volume is represented by the Maxwell-Boltzmann law:

$$N(v)dv = 4\pi N_T \left(\frac{m}{2\pi kT}\right)^{\frac{3}{2}} v^2 exp\left(-\frac{mv^2}{2kT}\right) dv \tag{2.1}$$

where N_T is the total number of molecules, m the molecular mass and N(v) dv is the number of molecules with speeds between v and v+dv.

2.1.2 Gyrofrequency, plasma frequency and Debye length

The gyrofrequency of ions and electrons are important parameters in governing the directions in which charged particles move in the ionosphere under collisions with their surrounding neutral species. Hence it is frequently used in calculations of important thermospheric parameters such as conductivities.

The gyrofrequency is a simple description of how quickly the charged particles move in a circular motion in the presence of a magnetic field as demonstrated by figure 2.1.

Each particle has a charge, e, velocity, **v**, and is under magnetic flux density, **B**. Hence the Lorentz force on the particle, **F**, which is normal to both e and **v** can be described by

$$\mathbf{F} = e.\mathbf{v} \times B. \tag{2.2}$$

The gyro-radius, r_b, of a particle of mass, m, is r_B=mv/Be, and the period of a revolution is P=$2\pi r_B$/v=2πm/Be. The particle is restricted to a circular path around the magnetic field, and the centripetal acceleration is provided by the Lorentz force. This leads to the angular frequency, or the gyrofrequency:

$$\omega_B = \frac{2\pi}{P} = \frac{Be}{m} \;\; rad\ s^{-1} \tag{2.3}$$

which is independent of the the velocity. For ions, i, and electrons, e, with the same kinetic energy, $\omega_i/\omega_e = m_e/m_i$.

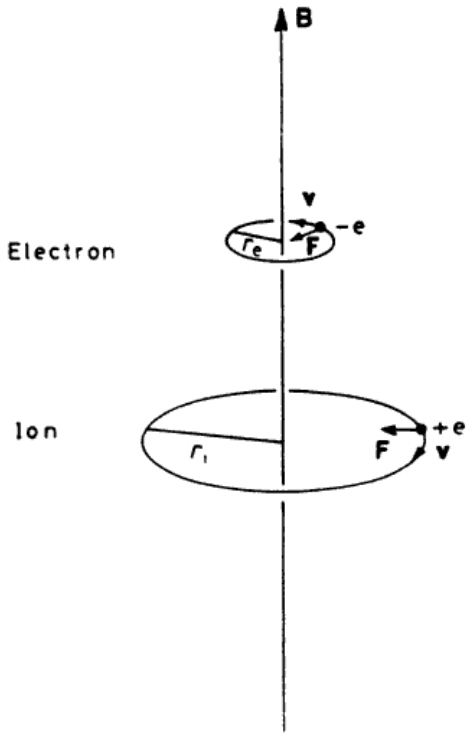


Figure 2.1: Courtesy Hargreaves [1992]. This figure shows an electron and a positive on gyrating around a magnetic field line.

Another main parameter in ionospheric studies is the plasma frequency. The plasma frequency is an important parameter for theories in radio propagation, which is therefore directly related to the radar data used in this thesis. Details of the radars can be found in section 2.4.3.

2.2 The neutral atmosphere

2.2.1 Vertical atmospheric structure

Figure 2.2 shows the altitude structure of the atmosphere in terms of temperature, composition and ionisation.

The atmosphere's layers and boundaries are determined according to the temperature gradients. The layers in order of height and according to temperatures are:

- Troposphere is the lowest, with temperature decreasing at approximately 10K/km with increasing height. The boundary of the troposphere is the tropopause, and is found at altitude of ∼12km.

- Stratosphere lies above, where the temperature increases with height, and reaches

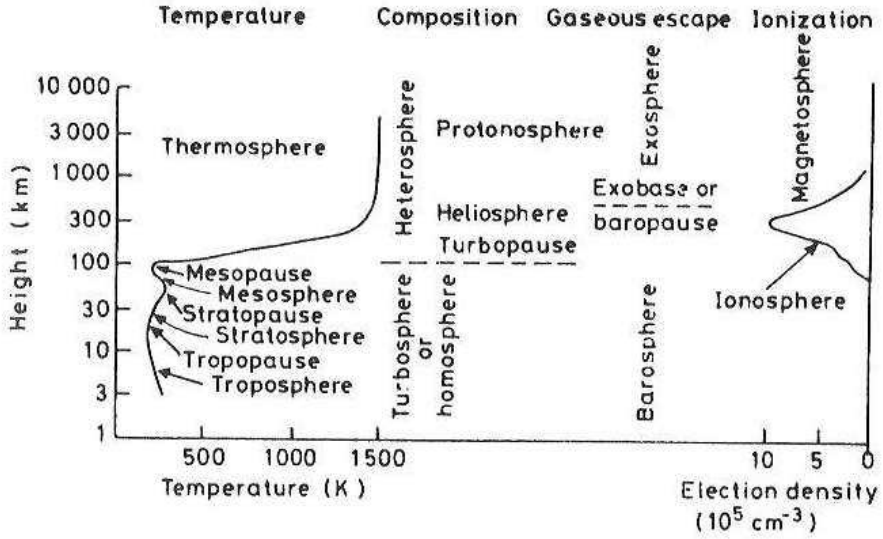


Figure 2.2: Courtesy Hargreaves [1992]. This is a summary of the neutral atmospheric altitude structure in terms of temperature, composition, gaseous escape and ionisation.

a maximum temperature i.e. the stratopause, at ∼50km altitude because of ultraviolet absorption due to its ozone content.

- Mesosphere is the next layer, where the temperature again decreases with height. This is the coldest layer of the atmosphere at ∼180K.

- Thermosphere is the hottest part of the atmosphere due to heating by solar ultraviolet rays, and its temperature quickly increases with height and reaches temperatures of ∼1000K. The thermosphere becomes mostly isothermal at altitudes above about 200km. This is also the region of the atmosphere that is observed by the FPIs.

2.2.2 Hydrostatic equilibrium

Hydrostatic equilibrium is fundamental to the physics controlling the 'barosphere', which is the atmospheric layer where gases are bound by gravity, in contrast to the exosphere. This is a region of the atmosphere where the density of gases decreases with height.

The scale height is the vertical distance where the pressure changes by a factor of e i.e. 2.718. It is larger for a hotter or lighter gas. The equation describing the scale height, H, is $\frac{kT}{mg}$. It is derived by combining the ideal gas law with the pressure difference over an element of height, dh.

A useful form of the hydrostatic equilibrium equation is derived with the substitution of the reduced height z=(h-h_0)/H. In this case, P=P_0 at h=h_0, and:

$$\frac{P}{P_0} = exp\left(-\frac{(h-h_0)}{H}\right) = e^{-z}. \tag{2.4}$$

The hydrostatic equilibrium can also be re-expressed in terms of number density, n, as follow:

$$n = n_0 exp(-h/H) \tag{2.5}$$

where n_0 is the number density at h=0. In reality, g, T and m are not constant with height. Furthermore, H is a simply a description of the variation of pressure with height, and has values of ∼5km at altitude of 80km and increases to as large as 70-80km at 500km altitude on Earth. This is caused by the decreasing mean molecular mass and increasing temperature with increasing altitude. The large scale height is partly responsible for ion outflow in the high latitude region.

2.2.3 Atmospheric temperature structure and heating

The heating and loss processes and their effect on neutral temperature are especially important for the FPI temperature trend study in chapter 4. First, the solar radiation spectrum is introduced in figure 2.3.

The Sun emits 4×10^{26}W of electromagnetic radiation and it irradiates Earth with almost 1.4kWm^{-2}. From this, about half of it reaches the surface of the Earth. For the interest of this thesis, the ionosphere is particularly affected by the short wavelength radiation: ultra-violet, extreme ultra-violet and the X-rays. These are absorbed at high altitudes and do not reach the Earth's surface.

The major production and loss processes that contribute to the shape of the temperature profile in the upper atmosphere are:

- **The absorption of solar ultra-violet and X-ray radiation.** The energy absorbed is consequently dissipated as heat from reactions such as photodissociation and ionisation and in general is the most important factor in the temperature structure. The basic mechanism is that when a photon is absorbed by a molecule/atom, the photon often has more energy than is used in the dissociation or ionisation processes, and the excess energy is therefore translated into kinetic energy of the gas and causes heating. Most radiation of wavelength less than 102.7nm (ionisation limit for O_2) is absorbed by the major upper atmospheric species of N_2, O_2 and O. Only

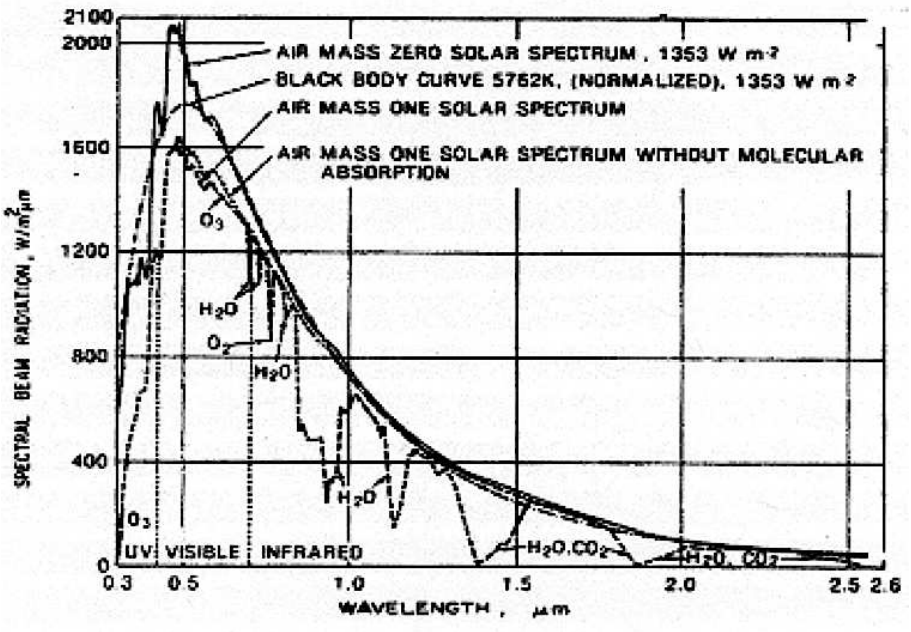


Figure 2.3: Courtesy http://www.fsec.ucf.edu/en/publications/html/FSEC-CR-670-00/. The figure shows the solar spectrum without any absorption outside of the atmosphere (air mass zero at 1353 Wm^{-2}), the black body curve at 5762K (at at 1353 Wm^{-2}) for comparison, the spectrum at the ground (air mass one) and the spectrum at the ground without molecular absorption.

approximately half of the energy is re-radiated causing atmospheric emissions, and the rest causes local heating. This is a major heating source in the thermosphere. Although only 3 $\mathrm{mW/m}^2$ is absorbed at solar minimum, it is still able to increase the temperature significantly at high altitudes because of the small air density. The heating rate is proportional to the gas concentration and is independent of height at the higher atmosphere.

- **Energetic charged particles entering the upper atmosphere from the magnetosphere.** This is a process that is the most important at high latitude regions, where energetic particles have the most direct access to the thermosphere. This heating could dominate during geomagnetic storms, and peaks in the E-region at 115-130km altitudes.

- **Joule heating by ionospheric electric currents.** This is another important process in the high latitude regions and is also a topic of interest in this thesis. Ion-neutral coupling is an important factor that needs to be included in Joule heating calculations, and the FPIs provide useful measurements of the neutral parameters. Joule heating is the greatest at the E-region.

- **Dissipation of tidal motions and gravity waves by turbulence and molecular viscosity.** This is another factor, but waves are not a focus of this thesis.

The main heating and loss processes, and how they are interlinked is summarised by figure 2.4 [Chandra and Sinha, 1973]. The heating processes are appropriate for generating the diurnal temperature pattern in most parts of the Earth. Chandra and Sinha [1973] derive the diurnal heating budget of the thermosphere by first discussing the relative importance of different heating sources that may affect the thermal structure of the thermosphere. Most of these heating sources originate from the ultra-violet and extreme ultra-violet solar radiations. An energy balance equation is then solved (e.g. equations 2.37 to 2.39).

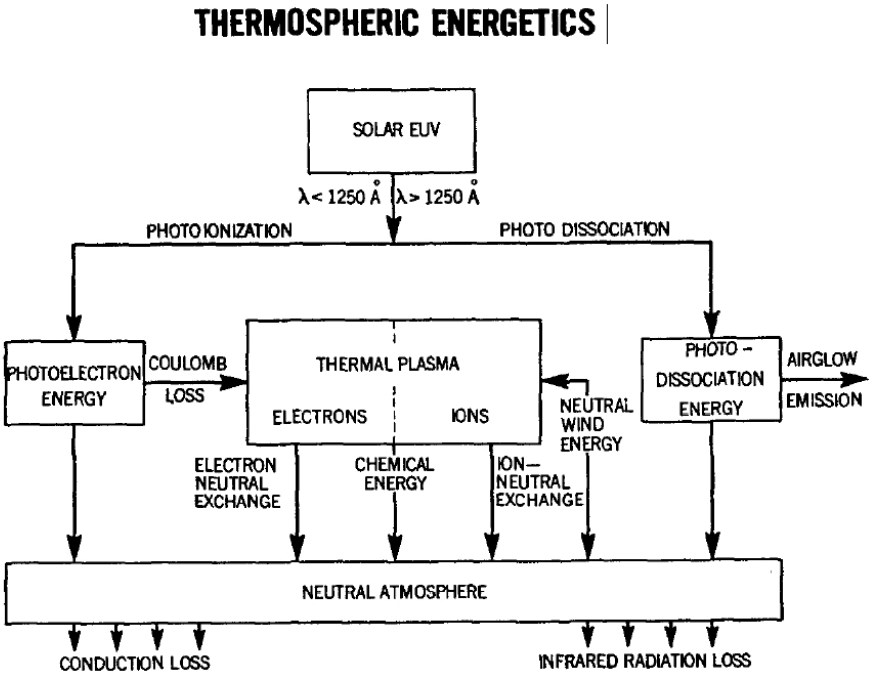


Figure 2.4: Courtesy Chandra and Sinha [1973]. The schematic represents the basic processes involved in thermospheric heating.

The height-integrated heating and loss rates for an unit cross section cylindrical column of gas at altitudes above 120km, over a typical 24 hour period from the Chandra and Sinha [1973] results are shown in figure 2.5.

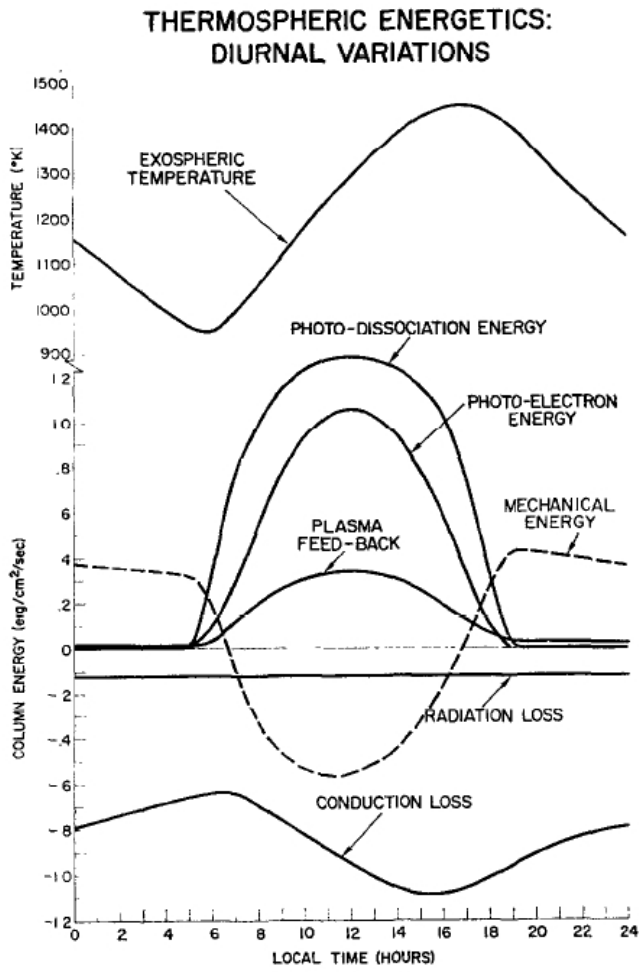


Figure 2.5: Courtesy Chandra and Sinha [1973] and quoted by Hargreaves [1992]. This shows the diurnal variation of altitude-integrated heating and loss processes for a vertical cylindrical column of gas with unit cross-section, for altitudes above 120km.

2.2.4 High latitude, night time neutral thermosphere heating sources and modelling

The average diurnal temperature pattern measured by SCANDI and the FPIs is investigated in section 4.3.7. The diurnal pattern presented in figure 2.5 is appropriate for the general global temperature. However, the night time, high latitude region requires additional heating sources, in order to correctly describe the diurnal pattern seen in the SCANDI/FPI data. It includes heating sources such as particle precipitation from aurora and cusp, while factors such as direct photoionisation are negligible since the optical instruments can only operate during the nighttime.

There are a number of models available, both empirical and numerical, that output neutral thermospheric temperatures, which includes e.g. Mass Spectrometer - Incoherent

Scatter (MSIS) [Hedin, 1987, 1991]; NCAR Thermospheric General Circulation Model (TGCM) [Roble and Ridley, 1987], and the UCL Coupled Mesosphere Middle Atmosphere-Thermosphere (CMAT2) model [Harris, 2001]. MSIS is used in this thesis to produce the long-term modelled neutral temperature for comparisons with FPI measurements, and the UCL CMAT2 model is used to investigate the relative importance of different heating terms in the observed diurnal pattern.

The thermospheric temperature variations are dependent on the balance between heating and loss processes, as well as transport mechanisms. For the high latitude regions, the following are the major heating and cooling processes considered by McCormac et al. [1988] from the TGCM model:

- Auroral Heat Input to a Parcel
- Cusp Heat Input to a Parcel
- Heat Input Due to Polar Drizzle
- Solar Heat Input to a Parcel
- Joule Heating of a Parcel
- Heat Loss Due to Downward Conduction and Eddy Diffusion
- Stability

The first three terms are all precipitation terms, which are calculated similarly by McCormac et al. [1988] with:

$$\Delta Q_0 = 35. \times 1.602 \times 10^{-12} \times Q_{total} \times \varepsilon(z) / \rho \tag{2.6}$$

where Q_{total} is the precipitation heating rate, $\varepsilon(z)$ is the altitude dependent neutral gas heating efficient described in Rees et al. [1983], ρ, is the neutral gas density, and the constant terms represent that 35 eV of energy per ion pair is released. The heating efficiency term, $\varepsilon(z)$ is dependent on the characteristic energy of the precipitating particles and the total energy flux.

The study in this thesis follows the work by McCormac et al. [1988], but the heating terms considered are appropriate for CMAT2 output [Harris, 2001], which are:

- Horizontal advection

- Vertical advection
- Radiation
- Horizontal conduction
- Ion-neutral interaction heating: including Joule and ion drag heating
- Vertical heat conduction
- Vertical viscosity
- Particle precipitation heating

The 'Particle precipitation heating' term in CMAT2 is equivalent to the 'Auroral Heat Input to a Parcel', 'Cusp Heat Input to a Parcel' and 'Heat Input Due to Polar Drizzle' terms in the TGCM model. The 'Joule heating' terms are specified separately in both models, and the remaining terms are defined differently in the two models, but the total output is expected to be equivalent. The neutral chemical heating and gravity wave drag heating are also considered in CMAT2, but their values are effectively zero at F-region altitudes.

The particle precipitation heating term is not readily output by CMAT2 at the time of this thesis, but a simple approach is taken here by applying equation 2.6, since both the total ionisation rate and neutral densities are available from CMAT2. In terms of the neutral heating efficiency, $\varepsilon(z)$, Rees et al. [1983] model the appropriate $\varepsilon(z)$ for particle precipitation phenomena (auroral, cusp or polar drizzle) using their respective characteristic energies and energy fluxes. An example of how the neutral heating efficiency changes with particle energies is shown in figure 2.6 [Rees et al., 1983]. For CMAT2, an average $\varepsilon(z)$ profile is used, irrespective of the type of particle precipitation. This is one of the factors that is under improvement for the model. Considering $\varepsilon(z)$ in figure 2.6, the fixed-value $\varepsilon(z)$ may introduce $\sim$10% error to the precipitation heating term, which is assumed to be negligible for the current version of CMAT2.

The behaviour of the heating terms is investigated to support the diurnal temperature trends measured by the FPI, with specific attention to the cusp influences, utilising the high spatial resolution from SCANDI.

In terms of the solar-cycle temperature trend, the neutral temperatures are expected to follow the solar activity represented by F10.7, as well as geomagnetic activity levels

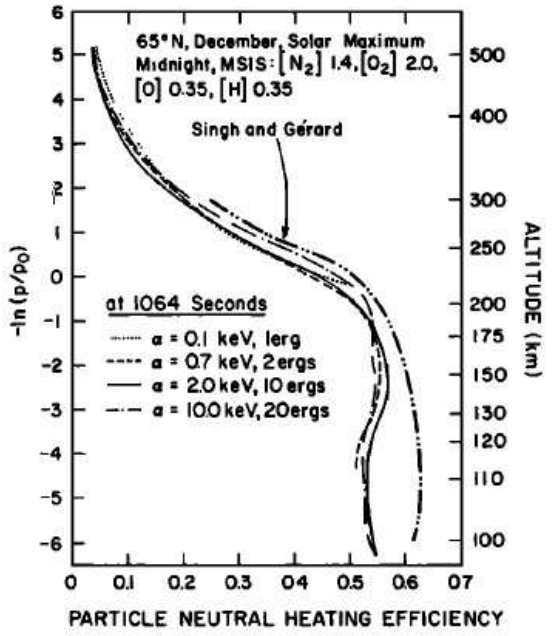


Figure 2.6: This is the generalised log pressure profile of particle neutral heating efficiency plot courtesy Rees et al. [1983]. The precipitation particles are auroral electron fluxes with Maxwellian energy spectra of several characteristic energies. The neutral atmosphere is appropriate for solar maximum conditions in December midnight, for 65^0N latitude.

represented by global Kp indices or the local K index. Figure 2.8 demonstrates the differences in thermospheric temperatures at solar maximum and minimum, for neutral (T_n), ion (T_i) and electron (T_e). However, for a high-latitude site directly affected by geomagnetic activities, maximum activity arrives ∼2-3 years after solar maximum. This is because sunspots appear at the solar poles and migrate towards the solar equator starting at solar maximum. Since the Earth is tilted at 6 degrees relative to the Sun's equatorial plane, the maximum activity is delayed by a few years from solar maximum. The delay is well demonstrated by over-laying the Ap indices and F10.7 values in figure 2.7. Figure 2.9 shows the average daily sunspot latitude locations.

There are previous studies that indicate a better correlation of neutral temperatures with solar activity F10.7 [Killeen et al., 1995] for Thule, which is a polar cap site (geomagnetic co-ordinates: 76.5N 68.0W). The Longyearbyen site studied in this thesis is located at a lower latitude, and is slightly closer to the auroral oval (geomagnetic co-ordinates: 75.4N 111.4E). The data set in this thesis is also longer: it covers 10 years instead of 6 years of data for the Thule study. Also, there are more recent studies by Aksnes et al. [2007], which show correlation of 200km altitude neutral temperatures with the auroral

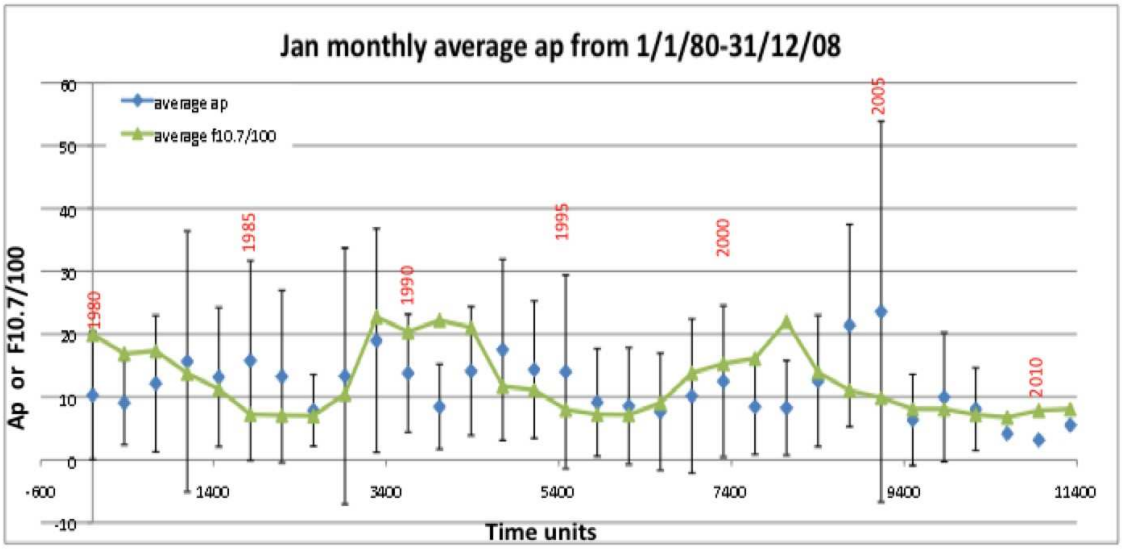


Figure 2.7: A plot of over-laid F10.7 and Ap values. Courtesy A. L. Aruliah, 2014.

electrojet index AL, with a 13-18.5 hours delay. The temperature trend study here includes correlation tests of FPI neutral temperature data with a number of activity indices, including F10.7, S10.7 [Dudok de Wit and Bruinsma, 2011, Tobiska et al., 2008, Floyd et al., 2002], AE (auroral electrojet indices) and Kp. Therefore it can demonstrate and suggest the best indices to be used as a proxy for atmospheric models such as CMAT2, for the high latitude auroral region.

2.2.5 Atmospheric circulation, winds and tides

The atmosphere can be treated as a fluid, and this assumes that there are a sufficient number of collisions between the gas particles within the typical time and spatial scales for the phenomenon of interest. This is a good assumption even at high altitudes of 400km, where a particle only makes one collision for every 3 hours. The fluid assumption is still valid because a typical scale of minutes and metres would still involve a large number of particles, 10^{13} m^{-3}. An extended version of Newton's second law of motion can be applied:

$$\frac{D\mathbf{U}}{Dt} + 2(\Omega \times \mathbf{U}) = -\frac{1}{\rho} \cdot \bigtriangledown p + \mathbf{g} + \mathbf{F} \tag{2.7}$$

where $\bigtriangledown p$ is the pressure gradient and ρ is again the density, $\mathbf{g}$ is the gravity, $\mathbf{F}$ is the acceleration due to other possible forces such as viscosity, and the resulting acceleration is described by

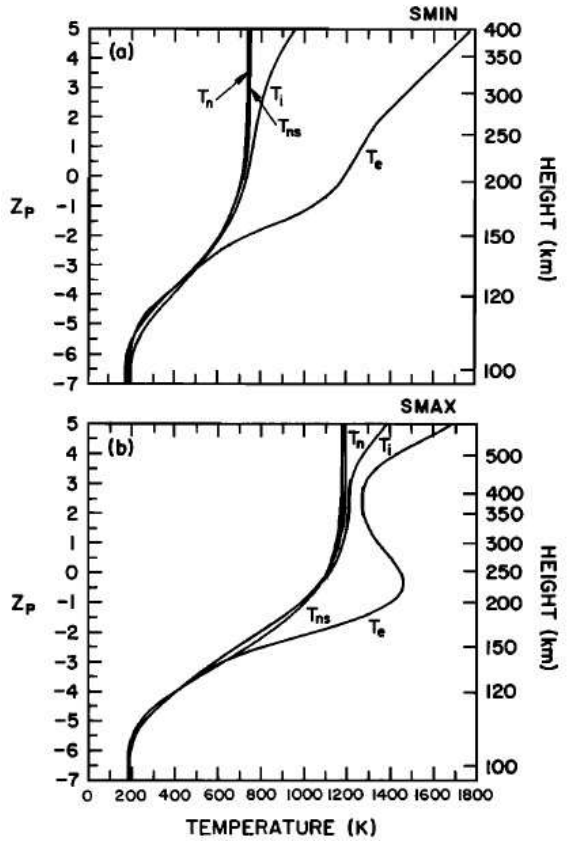


Figure 2.8: This figure shows the numerically-modelled global mean neutral (T_n), ion (T_i) and electron(T_e) temperature profiles at solar minimum (top) and maximum (bottom), courtesy Roble et al. [1987], Schunk and Nagy [2000d]. $T_{ns}$is the neutral temperature from the empirical model, MSIS.

$$-\frac{1}{\rho} \cdot \bigtriangledown p. \tag{2.8}$$

The operator is

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{U} \cdot \bigtriangledown \tag{2.9}$$

and it includes time variation of the fluid itself, $\frac{\partial}{\partial t}$, and due to advection, $\mathbf{U} \cdot \bigtriangledown$, where $\mathbf{U}$ is the velocity. Advection is the change seen by a stationary observer moving with the fluid. Finally, the second term in the equation is the Coriolis force, where Ω is the Earth's angular velocity. The Coriolis force is caused by rotation of the Earth.

It is generally assumed that the motions of interest are limited to the horizontal plane. This is because vertical motions require large potential energy changes due to the relatively large downward gravity.

For the upper thermosphere, which is the region of interest in this thesis, solar heating

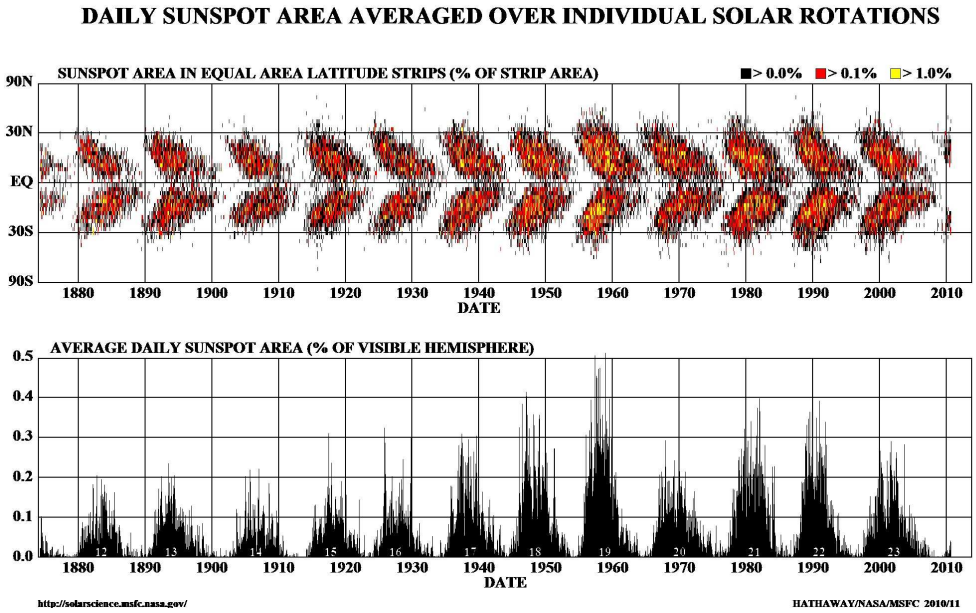


Figure 2.9: This is a butterfly diagram showing the migration of sunspots from solar poles to equator. Courtesy Hathaway and Adams [2010]

is one of the major heating sources and causes tides. The solar diurnal tide dominates above 250km and between 100-250km. Both diurnal and semi-diurnal components can be present.

The ion drag is treated as a retarding force on the thermosphere due to collisions between the neutrals and ions, in which the ions follow the geomagnetic field. The tidal equation (equation 2.7) can be reduced to

$$\frac{D\mathbf{U}}{Dt} = -\nu_{in}(\mathbf{U} - \mathbf{v}_i) - \bigtriangledown\Psi \tag{2.10}$$

where the ion drag dominates over the Coriolis force, and ν_{in} is the ion-neutral collision frequency, $\mathbf{v}_i$ is the ion velocity and $\bigtriangledown\Psi$ is the force producing the tide. Note that a more general approach to the tidal equation is to simply include the ion drag term described above, as well as the viscous drag force as part of the 'other forces' term, $\mathbf{F}$ in equation 2.7, which gives

$$2\mathbf{\Omega} \times \mathbf{U} + \frac{\partial \mathbf{U}}{\partial t} + (\mathbf{U}\bigtriangledown)\mathbf{U} = \mathbf{g} - \frac{1}{\rho} \bigtriangledown P - \nu_{in}(\mathbf{U} - \mathbf{v}_i) - \bigtriangledown\Psi + \left(\frac{\mu}{\rho}\right) \bigtriangledown^2 \mathbf{U} \tag{2.11}$$

where the final term is the viscous drag term, and μ is the coefficient of viscosity.

The neutral wind is smaller by day at about 40 m/s and larger at night at ∼120 m/s

because of the smaller ion drag at night. The thermospheric pressure gradients are directly affected by the temperature distribution, and have maxima and minima over the equator at 1500 and 0200 local time.

The amplitude of an upward propagating tide increases with altitude, similar to any upward propagating wave, to conserve energy density, $\rho A^2/2$, where A is the amplitude of the tide. Hence, as density decreases with height, the amplitude must increase. In terms of waves, the most important form of waves in the upper atmosphere is the acoustic-gravity wave. For a gravity wave, as a particle is displaced, the restoring force on the particle towards its original position is gravity. Gravity waves exist in the atmosphere because the density varies gradually with height. For 'normal' waves, the restoring force is usually the change in pressure because of the compression of the medium. In general both gravitational and compressional forces are taken into account.

For high-latitude, an additional heating source comes from energetic particle precipitation from the magnetosphere, as well as from the auroral electrojet. Note that these are significant sources of heating at high latitude, where the tides in general contribute 0.5 mW/m^2 at above 120km altitude, compared to the 0.5 W/m^2 from the electrojet and a slightly smaller amount from particle precipitations.

2.2.6 Atmospheric composition

The atmospheric composition from ground to 2000km at an exospheric temperature of 800K is shown in figure 2.10. Note that in terms of the altitudes of interest for the FPIs i.e. ∼110-280km, the three major neutral species are O, O_2 and N_2.

At altitudes above ∼105km in the thermosphere, the temperature increases rapidly with height due to absorption of extreme ultra-violet (EUV) part of the solar spectrum, also molecular diffusion (laminar flow) becomes more important than turbulent mixing (eddy diffusion). A major process in the thermosphere is the dissociation of O_2 into atomic oxygen, by photons of wavelengths lower than 175nm:

$$O_2 + h\nu \rightarrow O + O. \tag{2.12}$$

Atomic oxygen has a lower atomic mass than the molecular mass of O_2 and N_2. This part of the atmosphere is an isotherm, therefore the molecular mass of each species controls the density decrease here, affected by gravity. Hence, the O density falls off slower than O_2 and N_2, and O dominates at altitude above ∼150km. Above ∼400km, the lightest

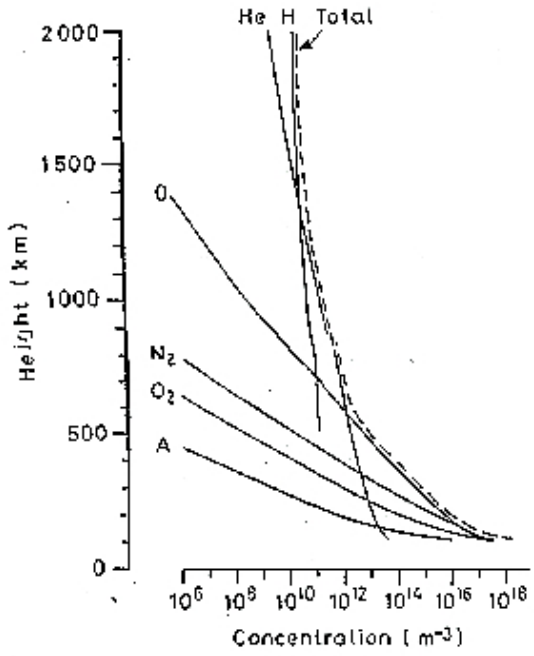


Figure 2.10: This figure shows the major gases in the thermosphere for an exospheric temperature of 800K. This are COSPAR International Reference Atmosphere data, 172.

species, Hydrogen and Helium become dominant. On average, the mean molecular mass is at 29 up to 105km, 25 at 200km and 19 at 400km (H=1, He=4, O=16, N_2=28, O_2=32).

The density of atomic oxygen plays an important role in the red and green-line emissions detected by the FPIs, which come from the excited metastable atomic oxygen atoms, O(^{1}D) and O(^{1}S). At the night-time thermosphere, the main source of these excited atomic oxygen are dissociative recombination of O_2^+ from thermal electrons.

2.3 High latitude thermosphere-ionosphere system

The high latitude ionosphere is one of the major topics in this thesis, and is directly coupled to the neutral thermosphere that the FPI measures. The research in this thesis investigates the atmosphere from the neutral thermosphere's perspective. This section provides an introduction to the coupled thermosphere-ionosphere system. Further information on the background of the ionosphere itself is provided in Appendix A.

2.3.1 Ionosphere: charged particle motion and conductivity

The charged particle motion and conductivity are vital for the understanding of the dynamics of both the charged and neutral thermosphere in the high latitude region.

There are two (extreme) types of charge motions in the atmosphere:

- Currents, when the positive and negative ions/particles move in opposite directions for the extreme case, and the ion and electron currents can be simply added together.
- Plasma drift, when the ions and electrons move with each other in the same direction, and there is no net current, but the bulk motion of the plasma remains.

The motions of the ions and electrons are influenced by different factors, including:

- Electric field, which is one of the driving forces that occurs naturally in the upper atmosphere.
- Ion-neutral collision, which is a complex interaction that describes collisions between the neutral wind/air and the ions. This requires the consideration of collision frequency (i.e. momentum transfer) between the particles and the gyrofrequency because of the geomagnetic field. The gyrofrequency is described by equation 2.3, and is $\omega_e = \frac{Be}{m_e}$ for an electron, and $\omega_i = \frac{Be}{m_i}$ for an ion. Note that the ions and electrons gyrate in opposite directions. The collision frequencies are important for determining the ionospheric conductivity.

The effect of collisions on charged particles is discussed first. Consider a charged particle in a magnetic field, and a driving force, F, acting along the magnetic field. At the same time, the charged particle collides with the neutral particles around it at frequency, ν. The drag force due to collision is therefore mvν and gives

$$v = \frac{F}{m\nu}. \tag{2.13}$$

Since F is in the same direction as the magnetic field, the magnetic field is not involved here. Next is to consider a more general case, where the particle motion can be described by balancing the different forces acting on the particle:

- The driving force, $\mathbf{F}$, in the plane perpendicular to the magnetic field
- The Lorentz force due to the geomagnetic field, $\pm e\mathbf{v} \times \mathbf{B}$, which is defined as positive for a position ion, and negative for an electron.
- The drag force due to collision, m$\mathbf{v}\nu$.

The co-ordinate system that is applied here is a 'right-hand' co-ordinate system, where the z-axis is parallel to the magnetic field, B. The force, **F**, is now considered to be parallel to the x-axis with magnitude, F_x. This leads to the force-balance equation:

$$\mathbf{F} \pm e\mathbf{v} \times \mathbf{B} - m\nu\mathbf{v} = 0. \tag{2.14}$$

The velocity components can be deduced using the substitution of the gyrofrequency, ω=eB/m; the fact that the second term, $\mathbf{v} \times \mathbf{B}$, acts in a direction perpendicular to the drag force, $m\nu\mathbf{v}$; and the substitutions, tanθ=$\pm$eB/mν and cosθ=mνv/F from the vector directions as described in figure 2.11.

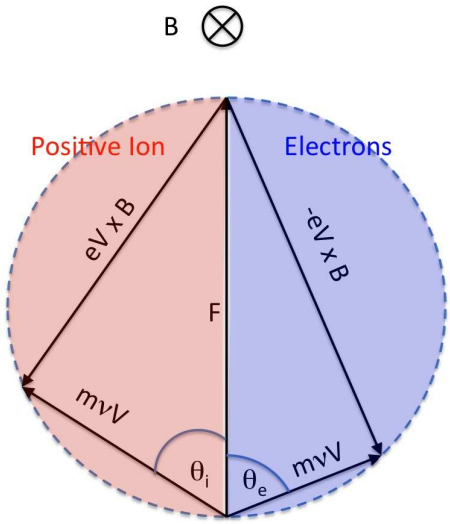


Figure 2.11: This figure is adapted from Hargreaves [1992], and shows the relationship between the different forces on positive ions (left) and electrons (right) under a magnetic field along z-axis, drag force, and force F (due to electric field. The diagram is in the x-y plane.

The velocity component results are shown in table 2.1. One of the important factors is the angle (θ_e or θ_i) between the forces from the external electric field and the particle velocity (see figure 2.11). The angle can simply be described by $f_0 = \frac{\omega}{\nu}$ in table 2.1, which is also tanθ=$\pm\frac{eB}{m\nu}$=$\frac{\omega}{\nu}$.

The ratio $\frac{\omega}{\nu}$ determines the ion and electron directions in the atmosphere, and determines if a part of the ionosphere has a current flow (electrons and ions in opposite direction: an extreme case) or a bulk plasma drift (electrons and ions in the same direction: also an extreme case).

Two main types of possible external forces need to be considered, which are

- A force due to wind in neutral air, F=mμU, where U is the neutral wind velocity. It

Table 2.1: This table shows the velocity components of an electron and positive ion in a magnetic field, under an external force and drag force due to collisions.

Velocity component	Electrons	Positive Ions
v_x	$\frac{F_x}{eB} \cdot f_1(e)$	$-\frac{F_x}{eB} \cdot f_1(i)$
v_y	$\frac{F_x}{eB} \cdot f_2(e)$	$-\frac{F_x}{eB} \cdot f_2(i)$
v_z	$\frac{F_x}{eB} \cdot f_0(e)$	$\frac{F_x}{eB} \cdot f_0(i)$
where	$f_1(e) = \frac{\omega_e \nu_e}{\nu_e^2+\omega_e^2}$	$f_1(i) = \frac{\omega_i \nu_i}{\nu_i^2+\omega_i^2}$
	$f_2(e) = \frac{\omega_e^2}{\nu_e^2+\omega_e^2}$	$f_2(i) = \frac{\omega_i^2}{\nu_i^2+\omega_i^2}$
	$f_0(e) = \frac{\omega_e}{\nu_e}$	$f_0(i) = \frac{\omega_i}{\nu_i}$

is assumed to be in a direction perpendicular to the magnetic field, which is a good approximation for high latitude site studies in this thesis. The neutral wind force can substitute F_x in the equations in table 2.1, and the neutral wind component along the magnetic field line can be estimated with Ucos I where I is the magnetic dip angle.

- An electric field, E_x, will induce a force eE_x for an ion, and $-eE_x$ for an electron, which can also substitute F_x in the equations in table 2.1.

Considering the above two forces, as well as applying the fact that at low altitudes, $\omega << \nu$, and for high altitudes $\omega >> \nu$, the electron and ion movements can be predicted for different altitudes, and it can be determined if the wind and electric forces cause a current or plasma drift at different parts of the ionosphere. The results are summarised in table 2.2.

The final parameter described in this section is the conductivity. The conductivity is simply

$$\sigma = \frac{i}{E} = \frac{Nev}{E} \tag{2.15}$$

where i is the current density in Am^{-2}, N is the density of the particles, e is the electric charge, and v is the particle velocity as derived. The derivations of the different components of conductivity i.e. Pederson conductivity, σ_1, and Hall conductivity, σ_2, are provided in Appendix B.

Table 2.2: This table summarises the effect of wind and electric field on the ions and electrons, and deduces if the two forces cause a current or a plasma drift at different ionospheric regions.

Altitude	Wind (U) force	Electric field (E) force	Additional Notes
High, above 200km F-region $\nu << \omega$ $f_0 >> 1$	Current $\perp$U Both e and i: $v_x << v_y$ e only: $v_y = \frac{m_e \nu_e U}{eB} = \frac{U}{f_0(e)}$ i only: $v_y = -\frac{U}{f_0(i)}$	Plasma drift $\perp$E Both e and i: $v_x << v_y$ Both e and i: $v_y = -\frac{E_x}{B}$	Small U current because of large f_0, E plasma drift =ExB drift dominates.
Middle E-region $f_0(e) > 1$ $f_0(i) < 1$	i current $\|$U Small e current $\perp$U	e current $\perp$E Small i current $\|$E	For f_0=1, i (75km) and e (120km) at 45^0 to U, e responds to E, i responds to U.
Low, below 80km $f_0 << 1$ $\omega << \nu$	Plasma drift $\|$U Both: $v_y << v_x$ Both: $v_x = \frac{m\nu U}{eB} f_0 = U$	Current $\|$E Both $v_y << v_x$ i only: $v_x = \frac{E_x}{B} f_0(i)$ e only: $v_x = -\frac{E_x}{B} f_0(e)$	Small E current, U plasma drift dominates.

2.3.2 Ion drag and Joule heating: theory and applications

The following is a summary of the derivations of ion drag and Joule heating based on Schunk and Nagy [2000e] and Zhu et al. [2005]. The contributions of neutral wind are often neglected in earlier studies because of the difficulty in directly measuring neutral wind in the thermosphere [Kelly, 1989]. This section provides both versions of the commonly used formulae involved in ion drag and Joule heating derivations, where the neutral wind contribution is either included or excluded. In order to determine the ion drag and Joule

heating from, for example, FPI and EISCAT data, the thermospheric conductivities and electric field have to be calculated first. These methods are well known and further details can be found in e.g. Schunk and Nagy [2000e] and Brekke et al. [1974].

In the strongly magnetised ionosphere, the current generated magnetic fields are too small to affect the large-scale plasma dynamics. So the intrinsic magnetic field, B, is taken as a known field. The currents can flow both along and across B in response to imposed electric fields, E. Both ions and electrons contribute to the current in the B perpendicular direction.

2.3.3 Perpendicular to B-field components

The perpendicular to B-field direction is considered, which is also the focus of this thesis because it has greater influence on the neutral thermosphere. The current in the cross-B direction is usually driven by a perpendicular E-field, $E_\perp$. The electric field dominates the perpendicular momentum equation (e.g. Schunk and Nagy [2000e] pg. 116) in the strongly magnetised ionosphere, and hence can be reduced to

$$\mathbf{E}_\perp = -\mathbf{u}_i \times \mathbf{B} + \sum_n \nu_{in} \frac{m_i}{e_i} (\mathbf{u}_i - \mathbf{u}_n) \tag{2.16}$$

where $\mathbf{u}_i$ is the ion velocity, $\mathbf{u}_n$ is the neutral wind, $\mathbf{B}$ is the magnetic field, e_i is the charge of the ion species, i, m_i is the average ion mass, and ν_{in} is the ion-neutral collision frequency with different neutral species, n. This equation is applicable to the ionosphere, where momentum transfer to electrons from ions is negligible because of the small electron mass, and also that momentum exchange between ion species is much smaller than with the neutrals, because $n_i << n_n$ and the ions drift at the same velocities. In some cases, for example, Brekke et al. [1974], a simplification is made on equation 2.16 that the ion-neutral collision frequency is very small in comparison with the ion gyrofrequency, Ω_i, and the collision term is ignored. The equation is reduced to

$$\mathbf{E}_\perp = -\mathbf{u}_i \times \mathbf{B}. \tag{2.17}$$

The altitude-dependent current density is given by

$$j = n_e e (\mathbf{v}_i - \mathbf{v}_e) \tag{2.18}$$

where

$$n_e = \sum_{i=1}^{n} n_i. \tag{2.19}$$

The perpendicular current density can then be expressed in terms of the conductivities and the E-field, and assuming there is no net charge in the atmosphere, as

$$\mathbf{J}_\perp = \sigma_P \mathbf{E}'_\perp + \sigma_H \mathbf{b} \times \mathbf{E}'_\perp \tag{2.20}$$

where $\mathbf{b}$ is the unit vector along $\mathbf{B}$, and $\mathbf{E}'_\perp$ is the effective electric field

$$\begin{aligned} \mathbf{E}'_\perp &= \mathbf{E}_\perp + \mathbf{u}_n \times \mathbf{B} \\ &= -(\mathbf{v}_i \times \mathbf{B}) + \mathbf{u}_n \times \mathbf{B} = -(\mathbf{v}_i - \mathbf{u}_n) \times \mathbf{B}. \end{aligned} \tag{2.21}$$

The Pederson (σ_P), and Hall conductivities (σ_H) are derived in section 2.3.1 as equations B.3 and B.5.

The electron contribution to the Pederson and Hall conductivities can be simplified because $\nu_e << \omega_e$ in most cases. Hence if only terms that are of order ν_e/ω_e are kept, and with the substitution $\Omega = e\mathbf{B}/m$, the conductivities are reduced to

$$\sigma_P = \sum_i \sigma_i \frac{\nu_i^2}{\nu_i^2 + \omega_i^2} = \frac{n_e e}{B} \frac{\omega_i \nu_i}{\nu_i^2 + \omega_i^2} \tag{2.22}$$

$$\sigma_H = -\sum_i \sigma_i \frac{\nu_i \omega_i}{\nu_i^2 + \omega_i^2} + \frac{\nu_e \sigma_e}{\omega_e} = -\frac{n_e e}{B} \frac{\nu_i^2}{\nu_i^2 + \omega_i^2}. \tag{2.23}$$

Ground-based magnetometers can only measure the height-integrated current, and the current density also has high variation with altitude. So for the purpose of correlating the ionospheric current measurements with ground-based magnetometers, the height integrated conductivity, and hence the integrated cross-$\mathbf{B}$ current density can be determined using the following assumptions:

- The neutral term, $\mathbf{u}_n \times \mathbf{B}$, in the effective E-field, $\mathbf{E}'$, is assumed to be altitude independent because the FPI can only measure the neutral wind at one altitude.
- The electrostatic E-field, $\mathbf{E}_\perp$, is also assumed to be height independent as it is in Brekke et al. [1974].
- It should be noted that during disturbed periods, the B-parallel components of the electric field and conductivity may have significant contribution to the current density, but they are excluded here.
- With the above considerations, equation 2.20 is integrated and gives:

$$\begin{aligned}\mathbf{J}_\perp &= \mathbf{E}_\perp \int_{z_2}^{z_1} \sigma_P dz + \mathbf{E}_\perp \times \mathbf{b} \int_{z_2}^{z_1} \sigma_H dz \\ &= \mathbf{E}_\perp \mathbf{\Sigma}_P + \mathbf{E}_\perp \times \mathbf{b} \mathbf{\Sigma}_H.\end{aligned} \tag{2.24}$$

It is interesting to see the directions of the current, and this can simply be determined with first expressing $\mathbf{E}_\perp$ as

$$\mathbf{E}_\perp = E_N \hat{e}_n + E_E \hat{e}_E \tag{2.25}$$

and the meridional and zonal components of the calculated currents are then

$$J_N = \Sigma_P E_N + \Sigma_H E_E \tag{2.26}$$

$$J_E = \Sigma_P E_E - \Sigma_H E_N \tag{2.27}$$

where N and E denotes the magnetic north (meridional) and east (zonal) directions respectively in the B-perpendicular plane, and $\hat{e}$ is the unit vector.

The local ion drag, $\mathbf{F}$, is $\mathbf{J} \times \mathbf{B}$, which is equal to $\mathbf{J}_\perp \times \mathbf{B}$ since $\mathbf{J}_\parallel \times \mathbf{B} = 0$. $\mathbf{J}_\perp$ is given by the local current density (not integrated) in equation 2.20.

The derivation of Joule heating can be found in Zhu et al. [2005] and Lu et al. [1995]. The energy exchange by Poynting's theorem gives

$$\mathbf{j}_\perp . \mathbf{E} = \underbrace{\mathbf{J}_\perp . \mathbf{E}^{'}}_{\varepsilon_J} + \underbrace{\mathbf{U}_n . (\mathbf{J} \times \mathbf{B})}_{\varepsilon_W} \tag{2.28}$$

which is applicable to a static case, and where the electric field component along B is much smaller than the perpendicular component, and is usually true for the ionosphere. The ε_J term is Joule heating, while the ε_W term is the rate of work done by the Ampere force on the neutral gases, the 'mechanical energy'. The total Joule heating rate can be extended to

$$\begin{aligned}\mathbf{J}_\perp . \mathbf{E}^{'} &= \sigma_P \left(\mathbf{E} + \mathbf{U}_n \times \mathbf{B}\right)^2 \\ &= \underbrace{\sigma_P E^2}_{\varepsilon_0} + \underbrace{\sigma_P \left|\mathbf{U}_n \times \mathbf{B}\right|^2 - 2\sigma_P \mathbf{U}_n . (\mathbf{E} \times \mathbf{B})}_{\varepsilon_1}.\end{aligned} \tag{2.29}$$

Here, ε_0 is the 'convection heating' i.e. the Joule heating rate in the absence of any wind, and ε_1 is the 'wind heating' which is dependent on both the neutral wind and the electric field/ion drift velocity. This term can be negative, if the neutral wind flows in the same direction as the ion drifts, since the the neutral wind speed is usually smaller than that of the ion drift at high latitudes.

2.3.4 The e-folding time

If an ionospheric circulation pattern is maintained for a sufficiently long amount of time, the neutral wind will follow the ionospheric pattern through ion-neutral interactions, since the plasma drift is at a higher velocity than the neutral wind due to e.g. the solar tide. Otherwise, the overall circulation will be determined by the balance between different ionospheric and neutral circulation forces. The important factor is the relative motion between the ions and neutrals, as described by the ion drag term, which is further described here (in additional to equation 2.10) as:

$$F = n_1 m \nu_{12} (v_2 - v_1) \tag{2.30}$$

where F describes the force on an unit volume of particles, 1 and 2 denotes two types of particles, which can be replaced with n and i. They are the neutrals and ions respectively. The equation can represent the ion drag force with $F_{ni} = n_n m \nu_{ni} (v_i - v_n)$, as well as the air drag force with $F_{in} = n_i m \nu_{in} (v_n - v_i)$, which are equal and opposite to each other, and leads to

$$n_n \nu_{ni} = n_i \nu_{in} \tag{2.31}$$

and hence the time constant:

$$\frac{1}{\nu_{ni}} = \frac{n_n}{n_i \nu_{in}}. \tag{2.32}$$

This describes the approximate time it takes for the neutrals to accelerate because of ion-neutral collisions. This is named the 'e-folding' time in Kosch et al. [2011], which applies high resolution ionospheric data from a new EISCAT radar scanning mode (section C), and the neutral data from SCANDI.

2.3.5 The general circulation in the high latitude thermosphere

The F-region ionospheric circulation is directly linked to the inner part of the magnetosphere and hence the solar wind. It can be explained in terms of open and closed magnetic flux tubes and plasma movements with the magnetic field lines, as demonstrated by figure 2.12.

It shows that the circulation pattern is formed by the following steps:

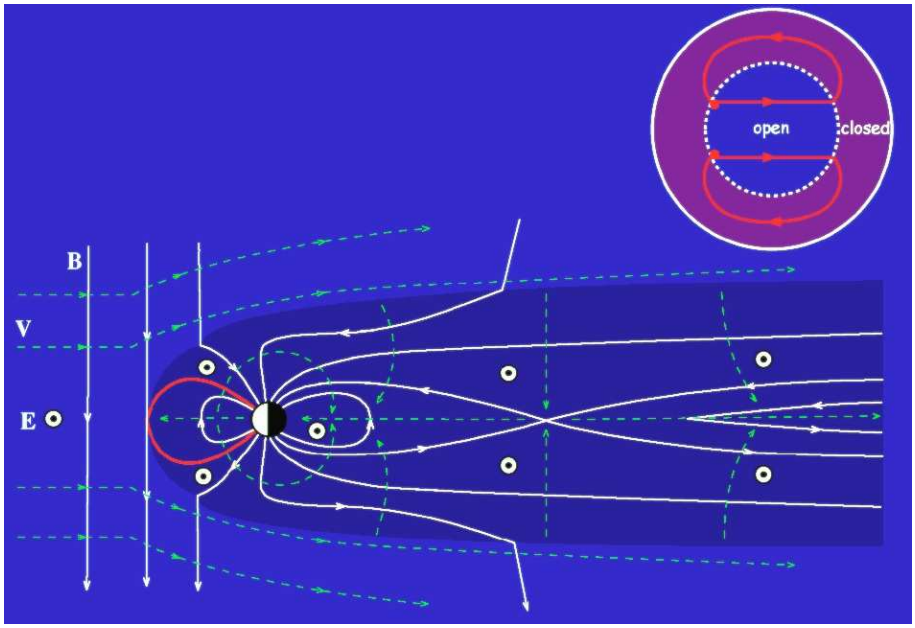


Figure 2.12: Courtesy Milan S. The magnetosphere (presentation), Leicester University. This figure shows the structure of the Earth's magnetosphere in the noon-midnight meridian plane. The circled-dot is the electric field associated with the plasma flow, the white lines are magnetic field lines, and the green lines are the plasma stream lines. The top right corner shows a view looking down on the northern high-latitude ionosphere, which shows the open and closed field line boundaries in white and the plasma flow stream line in red.

- As the plasma and magnetic 'frozen-in' condition is relaxed, the interplanetary magnetic field (IMF) is able to connect with the geomagnetic field lines at the magnetopause. A southward (negative) IMF is favourable for this connection (section 2.5.3). This causes 'open' magnetic flux with a magnetic tension force that accelerates the plasma, and these open tubes move towards the Earth's poles.

- These tubes are then moved by the magnetosheath flow to the tail.

- The open tubes eventually close again at the centre of the tail. The closed tubes then contract back toward Earth, and move back to the dayside at lower latitudes. This completes the cycle. The overall plasma flow pattern is shown in figure 2.13, and is known as the E×B plasma drift.

- The magnetospheric field lines are directly connected to the geomagnetic field lines in the ionosphere, which are approximately equipotentials and the magnetospheric circulation described here is mapped onto the F-region ionospheric plasma circulation.

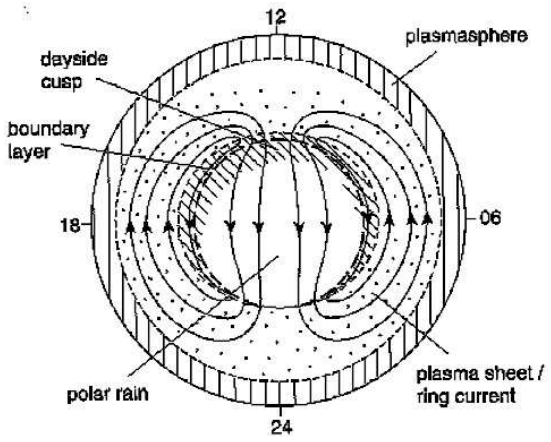


Figure 2.13: Courtesy Cowley S. W. H. The Earth's magnetosphere, http://www.ion.le.ac.uk/ education/ magnetosphere.html. This figure shows the general ionospheric flow in the F-region, and how the different parts of the magnetosphere are mapped onto the ionosphere.

The integral of the electric field gives the total electric potential across the polar cap, which has a maximum on the day side, and a minimum on the night side. The electric field can then be obtained from the gradient of the electric potential, and the velocity of the plasma is given by $\mathbf{v} = \frac{\mathbf{E}\times\mathbf{B}}{|\mathbf{B}|^2}$. Note that the electrostatic potential contours coincide with the streamlines of the plasma flow when there is only $\mathbf{E}\times\mathbf{B}$ drift. Hence, the electric potential is frequently used to show the general ionospheric circulation pattern as it is shown in figure 2.18.

In addition to the magnetospheric electric potential described above, the inner magnetosphere co-rotates with the Earth, which gives an additional co-rotational electric potential on the plasma drift. Hence, the overall pattern is not as uniform as the 'twin-cell' pattern shown in figure 2.13. Sources of electric potential and the relative influences of co-rotation and twin cell convention depending on latitude, need to be considered.

2.3.6 Cusp: introduction, review and application of FPI observations

The cusp is one of the main topics in this thesis and is introduced here. The definition of the cusp, which has been quoted by various authors such as Smith and Lockwood [1996], and taken from Newell and Meng [1988] is:

'The low-altitude cusp is the day-side region in which the entry of magnetosheath plasma to low altitudes is most direct. Entry into a region is considered more direct if more particles make it in (the number flux is higher) and if such particles maintain more of their original energy spectral characteristics.'

The Earth's magnetic cusps are an important area for study in space physics since they provide information on the magnetopause boundary and its variations. The cusp precipitation also has high influence on Joule heating rates and the ionospheric flow reaction/ inductive time to changes and transient events, and is known to show a dipole tilt effect [Newell and Meng, 1989] with seasonal and UT dependence.

There are numerous studies on the cusp with different instrumentation over the years, including ground-based observations with multichannel meridian-scanning photometers [Deehr et al., 1980], horizontal neutral wind studies with ground-based FPIs [Smith et al., 1985] and statistical studies of the cusp location from the Defence Meteorological Satellite Program (DMSP) satellite [Newell and Meng, 1988, 1989]. Smith and Lockwood [1996] provided a comprehensive summary on the cusp and related studies in 1996.

There is also renewed interest in the cusp due to recent instrumentation developments. For example, the CHAMP satellite which provides evidence of thermospheric up-welling [Lühr et al., 2004] and ion heating [Demars and Schunk, 2007] in the cusp region, at CHAMP altitudes of above 350km; the Streak mission which observes thermospheric cusp density [Clemmons et al., 2008]; the cusp-2002 sounding rocket which investigated ion upflow in the cusp [Burchill et al., 2010], as well as a statistical study of ion upflow and downflow from the Svalbard ESR EISCAT radar [Ogawa et al., 2009]. The most recent instrument is the RENU rocket, which was launched very recently on 12^{th}December 2010, and studied ion upflow in the cusp region with co-located ground-based instruments including the UCL FPI and SCANDI, EISCAT ESR radar, and other optical instruments.

2.3.7 Cusp precipitation energies and characteristics

It is well known that the cusps are typically located at about 78^o geomagnetic latitude, and are about 5^o wide. Their positions vary with factors such as IMF and geomagnetic activity [Deehr et al., 1980], as well as the solar cycle [Newell and Meng, 1989]. The magnetospheric plasma reaches the upper thermosphere with different energies and densities, depending on the topology of the magnetosphere. Hence different types of precipitation are mapped onto the high latitude thermosphere , including the cusp, and are shown in figure 2.14.

Please note 78^o is only a rough average. The cusp latitude from figure 2.16 are the average latitude over a specific year, from December 1983 to November 1984, which falls within the moderate activity level of a solar cycle. The cusp location is dependent on factors such as IMF and geomagnetic activity.

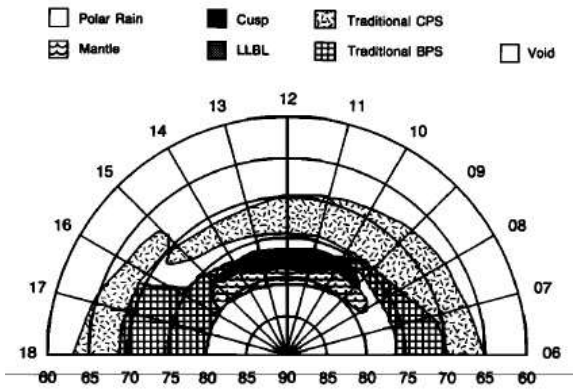


Figure 2.14: This is a schematic map showing regions where various types of precipitation are observed in the upper atmosphere. The plot is taken from Smith and Lockwood [1996], and it shows the results from a statistical survey of low-altitude satellite passes. It defines the magnetosheath precipitation including the cusp (the region of most ready plasma access), cleft (low-latitude boundary layer, LLBL, longitudinally extended region just interior to the magnetopause), mantle and polar rain (lower density regions); as well as the central plasma sheet (CPS) and boundary plasma sheet (BPS).

The 630.0nm and 557.7nm emissions can also be used to determine precipitation energy characteristics and hence the cusp location. The dayside aurora is thought to be related to the magnetospheric cusp or cleft and therefore the boundary layers of the magnetosphere. The results from a series of photometers are used by Deehr et al. [1980] to produce maps of red and green line emissions in relation to the cusp, and are shown in figure 2.15. The relative intensities of red and green-line emissions can give insight into the energy of the precipitation. Deehr et al. [1980] define the cusp or cleft region as a 630.0/557.7 intensity ratio of equals or exceeds 1. Their definition of a region of harder precipitation is where 630.0/557.7 equals one, and a transient event is when there is a short-lived enhancement of all emissions.

There are also numerous observations of when, in magnetic local time (MLT), the cusp is most likely to be detected. For example, Newell and Meng [1988] apply a year of DMSP satellite ion and electron precipitation data to locate the cusp (and cleft). They distinguish the cusp from other regions such as the boundary layer (cleft) by defining the cusp precipitations as those with the following characteristics:

- If the energy flux of the ions (electrons) is greater than 10^{10} ($6\text{x}10^{10}$) eV/cm^2 s sr. Energy lower than this value are neither cusp or boundary layer.

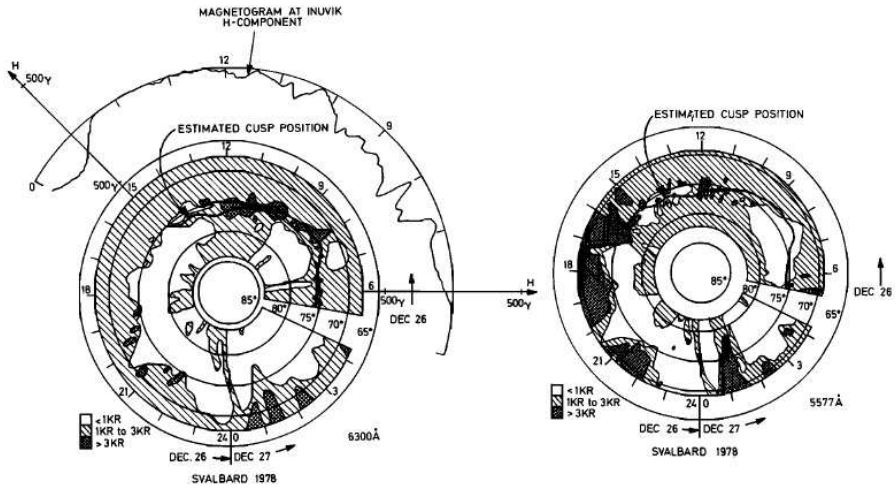


Figure 2.15: The left and right figures are the map of regions of different emission intensity in relation to the cusp, from red and green line emission respectively. They are taken from Deehr et al. [1980].

- If the energy flux in the 2- or 5-keV electron channel is smaller than 10^7 eV/cm^2 s sr eV. Since energy higher than this value indicates plasma sheet presence.
- Provided the above two points are true, the region detected is the cusp if the average ion energy, E_i, is between 300 and 3000eV, and the average electron energy, E_e, is less than 220eV.
- For boundary layer, again provided the first two points are true, the ion and electron energies are 3000eV $< E_i <$ 6000eV or 330eV $< E_e <$ 600eV.

The Newell and Meng [1988] results give the probability of detecting the cusp and the various other precipitations in terms of MLT. The variations of cusp latitude with MLT are also determined, and are shown in figure 2.16.

2.3.8 Cusp density structure

Clemmons et al. [2008], Demars and Schunk [2007], Lühr et al. [2004], Schlegel et al. [2005] are the cusp experiments mentioned so far. The first paper listed contains results from the Streak satellite mission, the second paper is a modelling paper trying to replicate CHAMP density measurements from the TGCM model, and the last two papers are CHAMP satellite data sets.

The thermospheric density structure reported by Streak and the CHAMP experiments varies. In general, CHAMP experiments, which measure higher altitudes of above 350km,

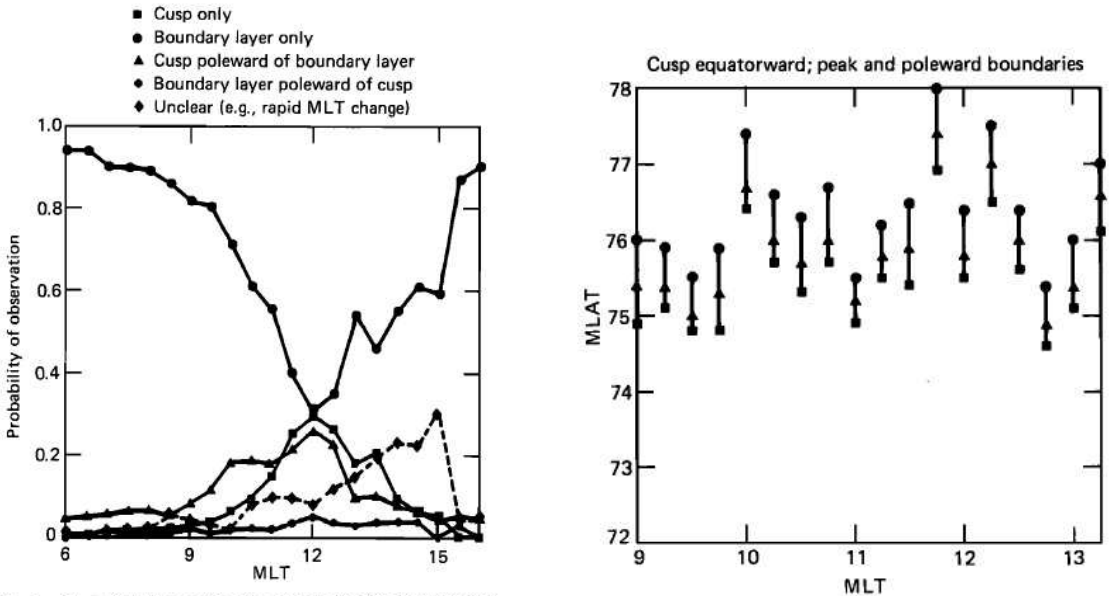


Figure 2.16: These are figures 2 and 5 taken from Newell and Meng [1988], the left plot shows the probability of observing the different precipitations vs. MLT, including those from the cusp. They are deduced from DMSP satellite data from 838km altitudes. The plot on the right shows the statistical cusp latitude vs. MLT. The cusp region is shown by heavy black lines.

report density enhancements in the cusp region [Lühr et al., 2004], and the larger CHAMP data set [Schlegel et al., 2005] also shows density minima at above 75^o magnetic latitude (MLAT), poleward of the density maxima on average at 75^o MLAT. The density maxima are usually observed between 10 to 11 MLT, while the minima can be found on both day and night sides. The magnitudes of the density changes, calculated from (density-density average)/density average, are 17.6% and -17.3% for the density maxima and minima respectively. The widths of the maxima and minima are on average 790km and 1070km respectively. These are Northern Hemisphere cusp results. The thermospheric density changes caused by the cusp are investigated in this thesis.

The Streak mission [Clemmons et al., 2008] measures lower altitudes between 123-325km, which is on average 150km lower than the CHAMP experiments, but is therefore a better match with the red and green-line emissions at ∼240km and ∼110km measured by SCANDI/FPI. The Streak results show that the density in the cusp is depleted relative to adjacent areas by 1-2%. The weak density depletion in the cusp contrasts with CHAMP's measurements of significant enhancment, and Clemmons et al. [2008] attributes this to the different altitude regimes of the two experiments.

2.3.9 Proposed cusp upwelling mechanisms

There are a number of potential physical mechanisms for forming the cusp density structures. Those most commonly discussed by various authors are Joule heating, particle precipitation, solar EUV radiation and dissipation of atmospheric gravity waves. The CHAMP-related analysis tends to favour the Joule heating mechanism, and recent studies have inferred that Joule heating rate should dominate [Schlegel et al., 2005]. It is also usually assumed that atmospheric gravity waves would have dissipated by the time they reach the upper thermosphere. EUV radiation and other large global processes are also ruled out by the TGCM model results used in Schlegel et al. [2005], Demars and Schunk [2007]. Hence Joule heating is considered to be an important source of the density changes observed in the cusp.

Schlegel et al. [2005] conclude that the Joule heating estimations are too small to sustain the upwelling and density enhancement observed by CHAMP, and propose that it is possibly because the neutral wind is neglected in their calculations. The effective electric field may have been twice as large if neutral wind effects were added. The model results of Demars and Schunk [2007] show that if the ion-neutral frictional heating is artificially increased in the cusp region, it can cause a neutral 'fountain', with upwelling of neutral gas in the heating region, and divergence and subsidence outside of the heating region at higher altitudes. A general increase in neutral temperature also occurs in and around the heating region, and highest temperature occurs near its poleward and equatorward boundaries.

However, the major mechanism suggested by Clemmons et al. [2008] from the Streak mission is particle precipitation, instead of Joule heating. This is because the Streak mission provides lower altitude measurements, in addition to the high altitude results from CHAMP. Therefore Clemmons et al. [2008] believe that direct heating by soft cusp and near-cusp precipitation can better explain the relatively weak density depletion in the cusp at lower altitude measured by Streak, as well as the high density enhancements measured by CHAMP. The idea is that particles generally lose most of their energies to collisions with neutral species before they reach altitudes below about 250km, and can heat the thermosphere above 250km sufficiently to produce the upwelling seen by CHAMP, but not at the Streak altitudes. To cause heating at lower altitudes, the precipitation needs to be relatively harder. The altitude effects shown by cusp experiments can be explained by the precipitation as a mechanism, but not Joule heating. The particle precipitation

model, as well as the neutral fountain Joule heating model are summarised in figure 2.17.

It is generally known that both Joule heating and particle precipitation heating play an important role in cusp heating and density change e.g. Carlson [2007].

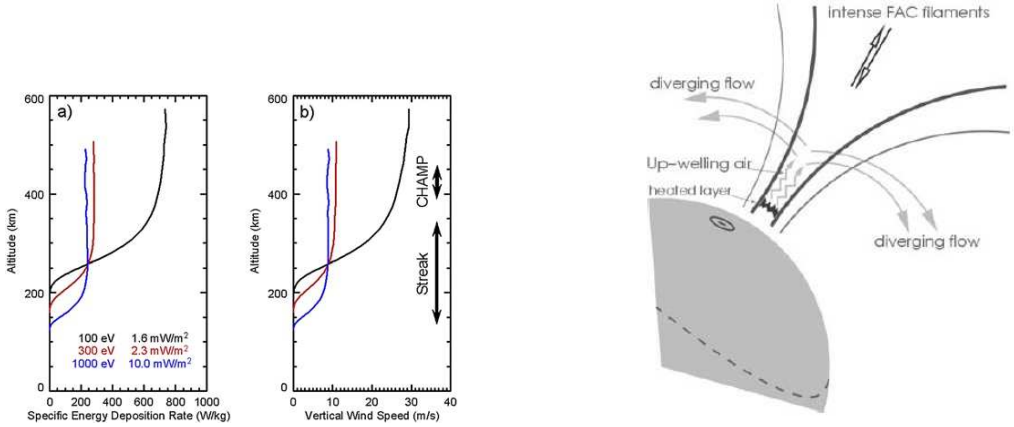


Figure 2.17: These are two mechanisms of density changes suggested by Clemmons et al. [2008] and Lühr et al. [2004]. The figures are taken from their respective papers. The figure on the left by Clemmons et al. [2008] is the modelling result of the profiles of specific energy deposition rate for precipitating electrons within the thermosphere for the left plot within the figure, the right plot shows the profiles of vertical wind speeds calculated from the energy deposition rates. The figure on the right by Lühr et al. [2004] is a schematic drawing of the thermospheric heating and up-welling suggested by CHAMP measurements.

2.4 Introduction to instrumentation

Data from a number of instruments are used in this thesis, in addition to the FPI and SCANDI data sets. This section gives an overview of these instruments. Details of SCANDI and the FPIs are not discussed here, but can be found in chapter 3.

2.4.1 Meridian-Scanning Photometer (MSP)

MSP data from the Longyearbyen, Svalbard Kjell Henriksen Observatory (KHO), provided by the University Centre in Svalbard (UNIS) and the Geophysical Institute, University of Alaska Fairbanks are used in this thesis.

The MSP detects atmospheric emissions with a cooled photo-multiplying detector. Narrow bandpass (50nm) filters at wavelengths of interest are mounted onto a tilting frame, which are placed in front of a rotating mirror that scans from north to south along

the magnetic meridian in the sky [Sigernes, 2002].

This MSP has 5 channels/filters, which includes the principle auroral wavelengths of 557.7nm, 630.0nm and 844.6nm from emission lines of atomic oxygen. The 437.8nm emission line from molecular nitrogen and the 486.1nm Doppler emission profile of hydrogen are also detected, and these two emissions are related to proton aurora.

Though SCANDI and the FPIs are also able to provide 557.7nm and 630.0nm emission measurements, they are non-calibrated and only the relative intensity variations are useful. The MSP are fully calibrated and hence provide the absolute value of the emission intensity in Rayleighs. Only the 557.7nm and 630.0nm MSP data are used in this thesis.

The MSP or photometers in general have a wide range of applications, some examples include: the monitoring of airglow and different forms/features/types of aurora; and spectroscopies of emissions providing insights into the energy deposition rate, the number flux and the characteristic energy of the bombarding electrons and protons. This can be achieved through modelling of chemical processes involved with different electron streams, under different atmospheric conditions and for different emission lines e.g. Rees and Jones [1973] and Rees and Luckey [1974]. For example, since 630.0nm and 557.7nm emissions are related to different electron precipitation with different energy spectra, the ratio between the measured intensities can give information about the characteristic electron energy. Direct bombardment by energetic electrons with energies below 1keV causes maximum ionisation in the F-region (i.e. red-line emissions), while E-region ionisation electrons have energies close to 3keV (i.e. green-line), and the ratio of these two types of electron energies can therefore be directly related to the ratio of red to green-line intensities that are measured by optical instruments.

2.4.2 Magnetometers

The most frequently used ground-based magnetometer data in this thesis are provided by the International Monitor for Auroral Geomagnetic Effects (IMAGE) network. Further information on the magnetometer network can be found in http://www.space.fmi.fi/image/.

One of the main foci of magnetometers is to study the auroral electrojets, where the IMAGE electrojet indicators adopt a similar method to the Auroral Electrojet AE indices i.e. for each time-step, the highest (IU) and lowest (IL) northern component of magnetic fields measured at selected stations are used to determine the average current, IE. The

aim is to estimate the total eastward and westward currents crossing the magnetometer network. Note that IU, IL and IE are used instead of the AU, AL and AE here because the indices measured here are only representative of the currents across the IMAGE network, which does not represent the actual auroral electrojet current provided by the A indices.

In terms of application in this thesis, the magnetometers provide the local variations in magnetic field in the form of geographic north, east and downward component of the magnetic field. It is essential for the calculations of electric field, current density, ion drag, conductivity, Joule heating etc., when combined with co-located measurements from the FPIs and the EISCAT radars, which provide the neutral and charged atmospheric parameters respectively. AE indices are also used in correlation tests with neutral temperature in the neutral temperature trend chapter.

2.4.3 Radars

The radars of interest in this thesis use the volume scattering technique, where electron scatter is detected when spatial variations/irregularities are present within a medium.

The separation between two scattering planes, d, is described by

$$d = \frac{\lambda}{2cos(\theta/2)} \tag{2.33}$$

where λ is the radar wavelength, and θ is the angle between the incident and scattered waves. This is because the signals scattered from irregularities that are spaced by $\lambda/2$ are reinforced in the direction back to the radar.

Volume scattering produces a continuous return, which has to be 'gated' to select the signal from each range. Gating every $10\text{x}10^{-6}$s gives a 'range gate' of (speed of light x $10\text{x}10^{-6}$ x 1/2)=1.5km, and the gate at $1\text{x}10^{-3}$s corresponds to a range of 150km [Hargreaves, 1992]. In general, the intensity of the return is related to the strength of the scattering mechanism, and frequency is related to the Doppler shift in the line-of-sight (LOS) or mirror directions. The shape of the spectrum can also contain information such as temperature of the medium.

The two types of radars that are frequently used in the high latitude atmosphere apply the coherent and incoherent scattering techniques. Coherence is basically a description of how quickly the measured medium changes (i.e. its stability) compared with the radar's ability to resolve the changes.

In general, 'coherent' radars usually take measurements from structures within the

medium that tend to vary more slowly relative to the radar's time resolution. Therefore, successive echoes can be added together, since there is little change in amplitude and phase between them. The Super Dual Auroral Radar Network (SuperDARN) is an example.

'Incoherent' radars usually take measurements from the thermal fluctuations of the medium with short coherence time i.e. the amplitude and phase of successive echos are unrelated, and can only be added together in terms of their power and autocorrelation. The European Incoherent SCATter (EISCAT) radar is an example. Data from the EISCAT radar are used frequently in this thesis, because they provide vital information on the ionosphere and are co-located with the SCANDI and FPI instruments. Some technical details of the EISCAT radar are provided in Appendix C.

There are many applications for the EISCAT radar, but this thesis focuses on ion-neutral coupling studies by the radar with co-located FPIs and SCANDI. The most recent published results, as an example, is a joint SCANDI-ESR experiment which observed meso-scale observations of Joule heating near an auroral arc and ion-neutral collision frequency from 2^{nd} February, 2010 [Kosch et al., 2011].

2.5 Other data and measurements

2.5.1 Solar Radio Flux F10.7

F10.7 is a measurement of the flux of 10.7cm (2.8 GHz) radio emissions, which is one of the indicators of general solar activity. Other indicators include the sunspot number and the counting of the number of solar flares.

F10.7 is routinely monitored and reported in flux units of 10^{-22}W/m^2-Hz. It is highly correlated with the sunspot number, R, and the total solar irradiance, D. The correlation of solar irradiances with sunspot number and F10.7, F, from average data sets can be expressed as [Hargreaves, 1992]:

$$S = 1366.82 + 7.71 \times 10^{-3} R \tag{2.34}$$

$$S = 1366.27 + 8.98 \times 10^{-3} F. \tag{2.35}$$

The flux has been measured since 1947, and represents a measure of diffuse, non-radiative heating of coronal plasma trapped by magnetic fields over active regions [Tapping, 1987], and is a good representation of the overall solar activity levels. Furthermore, F10.7

also tends to follow the changes in the solar ultraviolet that influence the Earth's upper atmosphere and ionosphere. The solar UV flux is a major driver of stratospheric chemistry, and increases in ionising radiation can significantly affect temperature and conductivity in the ionosphere. A major portion of the ionosphere is formed by photo-ionisation of the neutral atmospheric gases by solar extreme ultraviolet radiation with wavelengths less than 130nm. F10.7 is a reasonable proxy for solar UV/EUV flux and hence ionisation levels in the thermosphere.

Another method of monitoring the influence of solar activity on the thermosphere is through direct measurement of EUV radiation from the Mg II index, which is a dimensionless quantity measuring mid-ultraviolet solar activity at 280nm [Heath and Schlesinger, 1986, Lean et al., 2011], dating back to 1978 [Viereck et al., 2004]. Since EUV is directly related to thermospheric ionisation and chemistry, it may be a more appropriate correlation tool for thermospheric reactions. F10.7 is chosen as the solar activity index. This is because the models that are involved in the model-data comparison work here, e.g. MSIS in section 2.6.1, apply F10.7 instead of Mg II index as their input.

2.5.2 Solar EUV Monitor (SEM) and S10.7

There are recent studies which compare correlations between SEM (SOHO Solar EUV Monitor, SEM), Mg II and F10.7 measurements with solar x-rays and energetic protons detections on the SOHO satellite [Floyd et al., 2002]. They also investigate the most appropriate EUV proxy for use in general thermosphere modelling [Dudok de Wit and Bruinsma, 2011]; and developed new solar indices for use in thermospheric density modelling [Tobiska et al., 2008]. Dudok de Wit and Bruinsma [2011] conclude that the 25-34nm band as measured by SEM is systematically better than F10.7 and Mg II indices in thermospheric density modelling, from the comparison of 14 years of mean densities between 1997 and 2010. They deduce that solar EUV forcing in their model is defined better with the S10.7 index. S10.7 is determined as follow:

$$S_{10.7} = (-12.01) + (141.23)x(Soho\ SEM_{26-34}/Soho\ SEM_{26-34mean}). \tag{2.36}$$

Soho SEM_{26-34} is the daily integrated 26-34nm emission measured by SOHO Solar EUV Monitor (SEM), and Soho $SEM_{26-34mean}$ is the normalising factor calculated from the common time frame mean value, which is 1.9955x10^{10} photons cm^{-2} s^{-1}. The normalising value is then converted to Solar Flux Units through linear regression with F10.7

with reference to the common time frame. This method expresses the SEM values in the same units and scale as F10.7. This is named S10.7. The correlation between F10.7 and S10.7 with FPI-measured temperature are investigated in this thesis.

2.5.3 Interplanetary Magnetic Field (IMF)

The solar magnetic field is 'frozen' into the solar wind and carried into space, and forms the Interplanetary Magnetic Field, IMF. This occurs because the solar wind is a collisionless plasma where electric currents flow with little resistance. According to Faraday's Law of induction, it would take infinite energy to move field lines through the plasma, therefore inducing an infinite current. Consequently the field lines are frozen into the plasma. At a distance of few solar radii from the Sun, the solar wind velocity steadily increases as the plasma moves radially away from the Sun and the lower corona, and eventually becomes supersonic.

As the 'frozen' magnetic field is drawn outward by the radial solar wind, the slow rotation of the Sun (2.7×10^{-6} rad s^{-1}) bends the field lines into spirals, with a spiral angle of $\sim 43^0$ (at Earth) with respect to a line connecting the Sun and the Earth. Details of the IMF can be found in e.g. Schunk and Nagy [2000b].

The IMF is constantly monitored by satellite instruments on e.g. the Advanced Composition Explorer (ACE), which provides IMF orientations when the IMF comes into contact with the magnetopause, in terms of B_z (North-South) and B_y (East-West) etc, as well as the total strength. The IMF is provided by the ACE at a position slightly upstream of the magnetopause. It is well-known that a southward turning of the IMF leads to an increased rate of magnetic merging of the Earth's geographic field and the IMF at the day-side magnetopause. Southward IMF can partially cancel Earth's magnetic field at the point of contact at the magnetopause. For example, it can introduce newly opened field lines which can be convected across the polar cap and into the magnetotail on the night-side, increasing magnetic stress in the tail, and leading to a thinning of the plasma sheet and then to magnetic reconnection e.g. Schunk and Nagy [2000c].

The interest of this thesis lies with the influence of changes in IMF on the upper thermosphere and the cusp. There are many studies on this topic, since the ionospheric convection patterns are directly affected by the IMF clock/tilt-angle i.e. the angle of the IMF plane at the magnetopause in terms of B_z and B_y. The neutral thermosphere then responds to the changes in convection pattern through ion-neutral interactions such as ion-

drag. Some examples of IMF-ionosphere-neutral thermosphere studies include: Richmond et al. [2003], Chisham et al. [2000] and references therein. However, to summarise, the convection pattern is closest to a typical 'twin cell' pattern when B_y is small with a strong southward B_z, when the twin cells are approximately equal in size. The relative size of the twin cell changes and becomes less balanced with increasing B_y. Strong southward B_z also increases the overall size of the high latitude ionospheric convection pattern, which extends to lower latitudes.

Weimer [2005] models how the electrodynamics in the ionosphere changes with different IMF clock-angle, and considers ion-neutral interactions and the effect on Joule heating rates. Figure 2.18 shows the electric-potential changes at different IMF clock-angles taken from Weimer [2005], and figure 4.27 in chapter 4 shows the Joule heating variations.

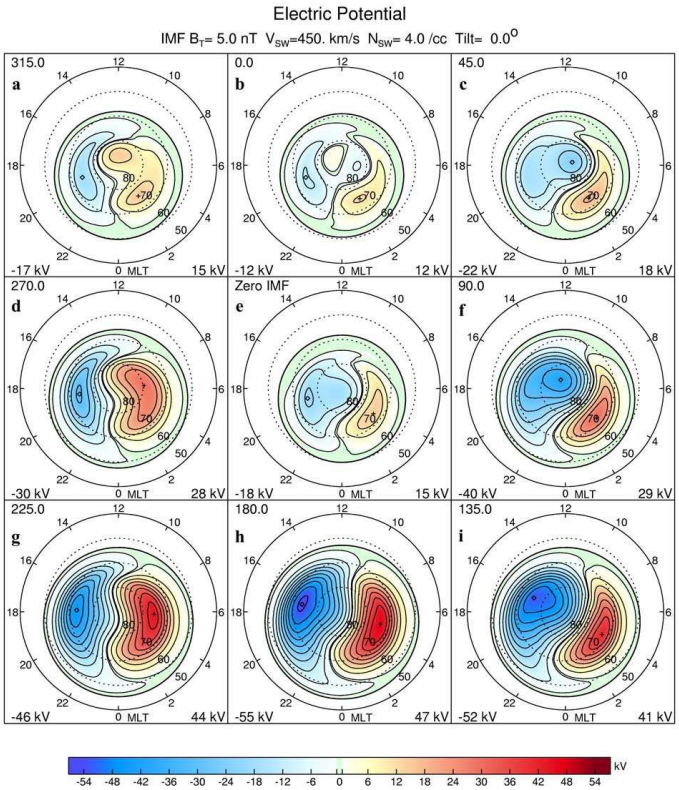


Figure 2.18: This figure shows the electric potential in the northern hemisphere mapped as function of latitude and magnetic local time (MLT), courtesy Weimer [2005].Each plot shows the electric potential predicted by Weimer [2005]'s model at different IMF tilt angle (at Bz-By plane), printed at the top-left corner of each plot. Minimum and maximum potential values are printed in the bottom left and right corners of each map. The middle plot shows results from tilt angle of 0^0. The IMF magnitude is fixed at 5nT and solar wind velocity at 450 km s^{-1} for all plots.

2.5.4 K, Kp and Ap Indices

The K index is a measure of geomagnetic disturbances, when the Earth's magnetic field responds to changes in Earth's current systems, in response to regular variations in solar radiation, interactions between the solar wind and the magnetosphere, and between the magnetosphere and ionosphere. The K index is simply a 'quasi-logarithmic' translation of readings from a single local magnetometer, and the magnetometers and their other uses are described in section 2.4.2. The K-index is a logarithmic local index, which records the geomagnetic activity in three-hour periods, relative to an assumed quiet-day curve. It consists of single-digit values between 0 and 9. The Kp index is also a 3-hour-range magnetic disturbance indicator. However, it represents the global average disturbance, deduced from the mean standardised K-index from 13 geomagnetic observatories between 44 and 60 degrees at northern and south latitudes for data provided by NOAA [2011]. The 3-hourly Ap index is the linear version (equivalent range) of the Kp index. The main use of the indices in this thesis is to identify if an experiment is performed under geomagnetically quiet, moderate or active conditions.

2.6 Thermospheric Models

Thermospheric models are essential tools in collecting measurements from different instruments and the numerous equations representing the physics involved. The models then output descriptions of thermospheric parameters under different solar and geomagnetic conditions. This section describes the models that are used in the thesis.

There are two basic types of models, numerical/mathematical models and empirical models. Numerical/mathematical models are usually composed of sets of equations, or computation from first principles, which are solved to obtain parameters of interest. The numerical models rely on the knowledge of the relevant processes involved for physically accurate output.

Empirical models simply recall parameters of interest from databases of accumulated data sets. Hence, they are a collection of data, which may have dependencies that can be represented by mathematical expressions. They involve no physics, and their accuracy is entirely dependent on their accumulated database.

Models can also be a mixture of empirical and numerical processes, which deduce some of the parameters mathematically, but some are taken from actual measurements in

a database.

2.6.1 MSIS

The Mass-Spectrometer-Incoherent-Scatter (MSIS) model is the most frequently used model in this thesis. Details of MSIS can be found in Picone et al. [2002], Hedin [1987, 1991]. The newest available version of MSIS is currently NRLMSISE-00 [Picone et al., 2002], and it is an empirical global atmospheric model that extends from ground to space.

MSIS input includes the time and date, altitude, latitude, longitude, local time, 81 day average F10.7 values, daily F10.7 from the previous day and daily magnetic index. The output includes number densities from major atmospheric species including O, O_2 and N_2, the total mass density and neutral temperature profile.

The NRLMSISE-00 empirical model is constructed with a database comprising the following measurements as described by Picone et al. [2002]:

- Total mass density from satellite accelerometers.
- Temperature from incoherent scatter radar covering 1981-1997.
- Molecular oxygen number density from solar UV occultation.

In addition, MSIS follows the concept that the thermosphere behaves similarly to a low pass filter, so spherical harmonics can be approximate eigenfunctions in the thermosphere e.g. Mayr et al. [1997], Mayr and Volland [1972], where some formulation of geographical and local time variations can be modelled based on low order spherical harmonics.

Both the densities and neutral temperature profile output are used extensively in this thesis.

2.6.2 CMAT2

The Coupled Middle Atmosphere Thermosphere (CMAT2) model is a three dimensional time dependent General Circulation Model (GCM) developed at UCL. It solves equations based on three laws of physics, which are:

- Conservation of mass:

$$\frac{\partial \rho}{\partial t} = \underbrace{-\overline{\nabla}_3 \overline{V}_3}_{3D\ velocity} \cdot \overbrace{\rho}^{mass\ density} \tag{2.37}$$

- Conservation of energy:

$$\overbrace{C_V}^{heat\ capacity} \cdot \frac{dT}{dt} + P\frac{dV}{dt} = 0 \tag{2.38}$$

- Conversation of momentum:

$$\frac{d\overline{V}_3}{dt} = \overbrace{-2\overline{Q} \times \overline{V}_3}^{coriolis\ term} + \underbrace{g}_{acc.\ due\ to\ gravity} - \overbrace{\frac{1}{\rho}\overline{\bigtriangledown}_3 P}^{pressure\ gradient} + \underbrace{\frac{1}{\rho}\overline{\bigtriangledown}_3 \left(u\overline{\bigtriangledown}_3\right)\overline{V}_3}_{viscosity} - \overbrace{\nu_{in}\left(\overline{V}_3 - \overline{U}_{i3}\right)}^{ion\ drag}. \tag{2.39}$$

Currently, the normal mode for CMAT2 is a grid of latitude, longitude and logarithmic pressure divided with 2^0 latitude, 18^0 longitude and 63 different pressure levels. Details of CMAT2 can be found in Harris [2001].

In terms of this thesis, CMAT2 provides the relative importance of different heating terms in the FPI-measured F-region neutral temperature diurnal pattern. Other outputs from CMAT2 include global neutral gas temperatures and wind velocities.

Another example of numerical model that is mentioned, but not applied in this thesis is the NCAR Thermospheric General Circulation Model (TGCM) [Roble and Ridley, 1987], which similarly solves the coupled continuity, momentum and energy equations for the thermosphere in 3-dimensions.

2.6.3 IRI

The International Reference Ionosphere (IRI) model is an empirical model that for given location, time and date, provides the monthly averages from 50 to 200km altitudes, of ionospheric parameters including: electron density, electron temperature, ion temperature, ion composition, ion drift; and ionospheric electron content (TEC) i.e. the total number of electrons present along a path between two points, where 10^{16} electrons/m^2=1 TEC unit.

The following instruments are some of the major data sources for IRI: the worldwide network of ionosondes, incoherent scatter radars, topside sounders and in situ instruments on several satellite and rockets. The information here are based on [NASA GSFC and Bilitza [2010]], and other references include Rawer et al. [1978], Bilitza et al. [1993].

The main application of IRI in this thesis is to provide the average hmF2 variation over the recent solar cycle, which is the height maximum of the F2 layer, where the ionospheric electron density reaches a maximum. The hmF2 is closely related to the peak red-line emission altitude, which is important in determining the altitude at which the FPI observes.

2.6.4 OVATION

The (Auroral) Oval Variation, Assessment, Tracking, Intensity, and Online Nowcasting, OVATION, model is an empirical model that provides an estimation of the auroral oval size, position and intensity of precipitation [OVATION, 2011]. It can provide historical data, as well as real-time nowcasting. The major data sources for OVATION are DMSP (Defence Meteorological Satellite Program) particle precipitation data; UVI, which is an imager on the NASA satellite, Polar, that takes global pictures of the aurora in ultraviolet wavelengths; MSP (Meridian Scanning Photometer), and SuperDARN data. Other OVATION references include Newell et al. [2002, 2010].

The main application of the OVATION model in this thesis is to estimate the locations of the auroral oval's poleward and equatorward boundaries.

Chapter 3

SCANning Doppler Imager, SCANDI

This is an instrumentation chapter which describes the new SCANning Doppler Imager, SCANDI. It was first deployed at the Auroral Station at Longyearbyen, Svalbard (78^oN 16^oE) in December, 2006, and was subsequently moved to the Kjell Henriksen Observatory (KHO) in November 2007.

SCANDI is an all-sky version of the original network of UCL Fabry-Perot Interferometers (FPI). For the red-line (630nm) emission observation mode, SCANDI's 150^o field-of-view (FOV) covers a circular area of radius at approximately 600km at 240km altitude in the F-region, which is the assumed altitude of the red-line emission. SCANDI is also capable of observing the green-line (557.7nm) emission, which occurs at 110km altitude and is representative of the E-region. SCANDI's FOV is divided into 25 observing zones as of observing season 2011, with horizontal spatial resolution of ∼100-300km. Hence, all-sky maps of neutral dynamics, wind and temperatures, as well as intensity are produced. Together with SCANDI, the UCL FPI network now consists of SCANDI and a FPI in Longyearbyen, Svalbard (78^oN 16^oE), which alternates between red and green line observations; one FPI in Sodankylä, Finland (67^oN 27^oE) observing red line; and two FPIs in Kiruna, Sweden (68^oN 20^oE), with one observing red line and the other green line. A summary of the UCL FPI network can be found in McWhirter et al. [2005]. In addition, colour all-sky cameras are installed at each site to provide continuous monitoring of sky condition and auroral structure information during geomagnetic events.

The FPIs are also conveniently located close to other thermosphere/ionosphere ground-based instruments. For the Svalbard site, the FOV of the UCL instruments overlap with

the EISCAT ESR radar, as well as the FOV of CUTLASS SuperDARN radars. The Kiruna and Sodankylä FPIs have a common observing volume, and can provide bistatic measurements. Together with the Nagoya STELab FPI in Tromsø, tristatic observations can also be made available, and the mainland EISCAT radar system can be organised to observe the same tristatic/bistatic volume. These allow simultaneous observations of the ionosphere and neutral thermosphere at both the polar cap and auroral oval regions (figure 3.1). Furthermore, the IMAGE magnetometer network provides valuable real-time observations of magnetic field. The co-located FOVs of SCANDI, the EISCAT ESR and the SuperDARN CUTLASS radars are shown in figure 3.1

This thesis chapter focuses on SCANDI's data analysis and calibration, as well as providing a summary of the instrumentation. SCANDI is mostly based on the work of Conde and Smith [1997], who constructed the Scanning Doppler Imager (SDI), which was installed in Poker Flat, Alaska. SCANDI's data analysis software is also based on the work by e.g. Conde [2002].

3.1 Division of responsibilities

SCANDI was built and is maintained by the instrumentation team of the Atmospheric Physics Laboratory (APL), UCL. The optical, mechanical and electronic design and construction was carried out by Dr. Ian McWhirter and Dr. Andy Charalambous. Dr. Ian McWhirter also wrote the instrumentation control and data collection software, and gives continuous support on the day-to-day running of SCANDI and the FPIs. The capacitance-stabilised etalon, which is an essential component of SCANDI, as well as background FPI theories can be found in McWhirter [1993].

In terms of data analysis, the LOS wind and wind vector analysis for the original FPI has been developed and performed by Dr. Anasuya Aruliah. Dr. Eoghan Griffin and the author are both involved in the adaptation of the original SDI data analysis software for SCANDI use. The SCANDI analysis method was subsequently also applied to narrow-field FPI analysis for neutral temperature data.

The author's responsibility includes the adaptation of wind vector analysis from Conde [2002]; the all-sky mapping/projection of SCANDI data from the fish-eye lens; the calibrations of the intensity, temperature and wind data before the final output, which includes testing for statistical variations and SCANDI-FPI comparisons; and developing different data display methods for scientific uses.

The author has also been directly involved in several field campaigns in Svalbard, Sweden and Finland. She has assisted in the assembly, alignment and calibration of SCANDI at the Adventdalen Observatory and subsequently the new Kjell Henriksen Observatory, both at Longyearbyen, Svalbard. Also she has been involved in the maintenance and calibration of the Fabry Perot interferometers at all three field stations.

Part of the calibration work discussed here is published in sections 2.3 and 2.4 in Aruliah et al. [2010]. Both sections are the original work of the author, and this chapter expands further on the work described in Aruliah et al. [2010], including additional input from recent developments. This chapter consists of original work by the author unless otherwise stated.

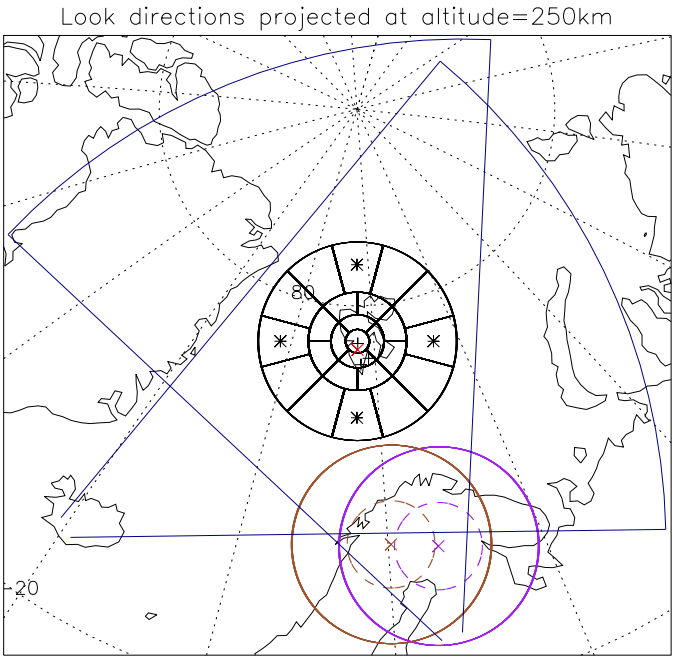


Figure 3.1: This figure shows the different instruments' FOVs at 250km altitude, and the centre island on the plot is Svalbard. The black circular grid is the 25 SCANDI zones in the 2006/2007 setting. The red cross in the middle is the field-aligned direction from the EISCAT radar, and the 3 black crosses are the 3 look directions of the EISCAT radar in CP-2 mode. The 4 asterisks within the outer SCANDI zones are the north, east, south, west positions of the FPI. The dark-blue lines define the FOVs of the CUTLASS radars. The observing positions from the Kiruna and Sodankylä FPIs are covered by the brown and purple circles respectively, where the solid circles are the F-region 250km altitude projections, and the dotted-line circles are the E-region, 110km projections.

3.2 SCANDI instrument

Early FPIs for upper atmospheric studies at high latitudes used photometers such as detectors, which requires integration times at the limit of about 5 minutes. Imaging Photon Detectors with good signal-to-noise ratio were then developed at the APL in the 1980s, which could provide two-dimensional array imaging, and reduced the integration time to 60s. Finally, with the newest development of Electron Multiplying Charge Coupled Device (EMCCD), high resolution two-dimensional image arrays with high-speed readout and good signal-to-noise ratio became available, and enabled the refinement of all-sky imaging FPIs such as SCANDI. Further history of FPI developments can be found in e.g. McWhirter [1993], Aruliah et al. [2010] and references therein.

Comparing SCANDI against the FPI, the most obvious difference is their FOVs. The FPI has a 1^o FOV that uses a scanning mirror, which rotates through its look-directions in a typical sequence of North, South, East, West, Zenith and calibration lamp. For SCANDI, with the EMCCD, a capacitance stabilised scanning etalon [McWhirter, 1993] and the replacement of the scanning mirror with a fish-eye lens, it is able to provide an instantaneous all-sky maps of neutral winds, temperature and emission intensities within its 150^o FOV.

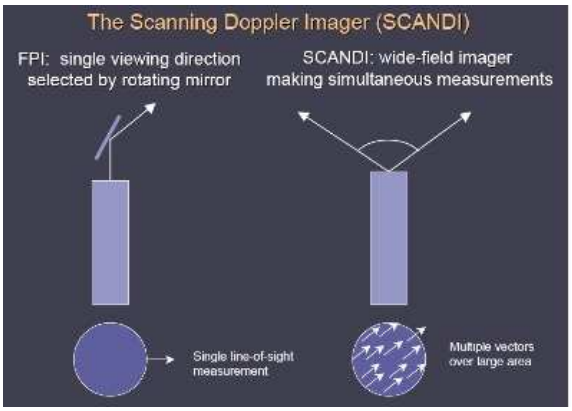


Figure 3.2: Courtesy Aruliah, A.. This figure contrasts the main differences between SCANDI and the FPI.

In order to determine the horizontal wind vector, the FPI needs to combine the line-of-sight (LOS) measurements from its rotating-mirror sequence ($\sim 60^o$ zenith). Each mirror position observation takes about 45s, including mechanical movement time. One complete sequence therefore requires $\sim$5 minutes. Hence, the final horizontal wind vector from the FPI is representative of the average wind field, over the circle determined by the observing

positions and the time taken to complete the scan. For SCANDI, horizontal wind vectors can be mapped for each of the 25 zones within its FOV. The relative emission intensity and neutral temperature are directly measured at each zone. As of November 2011, each SCANDI map requires 7 minutes to be produced, with a spatial resolution is much higher than those of the FPI. Figure 3.2 indicates the single line of sight FPI observation compared with a SCANDI wind field.

3.2.1 SCANDI hardware - Design and Optics

This section mostly follows section 2.1 of Aruliah et al. [2010], which is original work by the author. Parts of McWhirter [1993]'s thesis, including the description of the capacitance stabilised etalon (CSE), also provide vital background to the technical descriptions. Further detail on SCANDI's instrumentation hardware can be found in these documents.

A three-dimensional schematic diagram, as well as its simplified optics are shown in figure 3.3. The SCANDI instrument is three metres long and a cylindrical construction is used for most of the optical path. In order to maintain image stability, the entire optical system is temperature controlled. The left diagram in figure 3.3 shows the essential components of SCANDI, which also presents the adjustable floor-standing frame for height adjustments. The frame is important for the installation of the instrument at different locations. For example, the fish-eye lens at the top of the instrument is required to be at the optical centre of the observatory's dome to minimise dome reflections. This can be achieved by adjusting the height frame.

The maximum mechanical stability is created by first modelling the optics of the entire instrument by computer, which fixes the positions of all the optical components. The focal distances are then fine-tuned in the finished instrument with minimal adjustments.

SCANDI's simplified optics, including a ray diagram, are shown in the right diagram in figure 3.3, and the listed optical components are further described below.

3.2.1.1 The all-sky imaging lens

SCANDI uses a commercial equi-solid angle projection fisheye lens (sigma, f=8mm). The radial mapping of the fisheye lens is typically described by

$$r = \alpha * sin(\beta\theta) \tag{3.1}$$

where r is the distance from the centre of the camera image to the point of interest, α

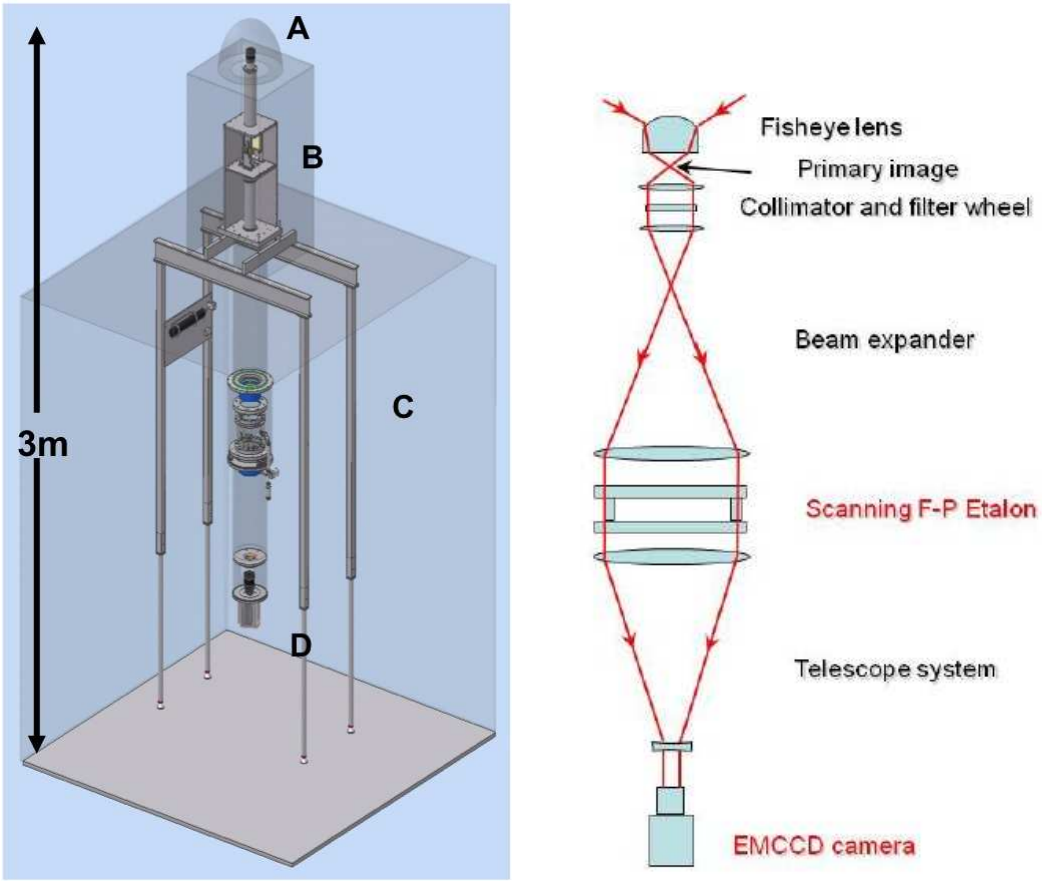


Figure 3.3: This figure shows the three-dimensional (left) and two-dimensional (right) schematics of SCANDI (courtesy McWhirter, I., Charalambous, A. and Aruliah et al. [2010]). SCANDI is about 3 metres long, and for the left diagram, the labels are: A. All sky Sigma 8mm Fish-eye lens, B. Laser and Calibration lamp panel and filter wheel, C. Capacitance-stabilised scanning etalon, D. EMCCD, Andor iXon camera. The diagram on the right shows the simplified optics (not-to-scale).

is a scale factor for converting from angle in space to a length in the image plane, β is the radial mapping parameter, and θ is the angle between a incoming ray and the optical axis. For a 'perfect' lens, α=2f, where f is the lens' focal length, and β=0.5. Figure 3.4 demonstrates the equi-solid angle fisheye lens projection for SCANDI. The theoretical fisheye lens equation here can introduce a 5-10% uncertainty to the projection, when compared to the experimental fisheye lens relation provided by equation 3.30 (section 3.7.1).

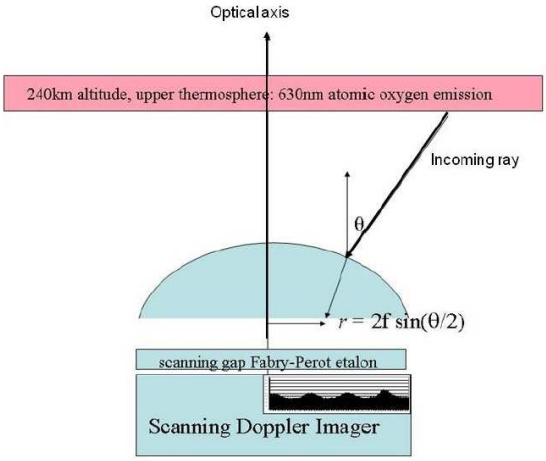


Figure 3.4: Courtesy Aruliah et al. [2010]. This figure shows the equi-solid angle fish-eye lens projection in SCANDI.

3.2.1.2 Filter, collimator and calibration unit

The collimator directs the incoming light from the fisheye lens into a 50mm diameter light path, which also contains the filters of the system. An interference filter is essential in order to reduce the amount of continuum in the observed light. The aim is to produce a level of light that gives a clear spectral line, but does not overload the detector.

Three filters are currently installed in the four-position filter wheel, with one of the positions left empty to allow broad-band imaging. The broad-band image is a white-light image of the fisheye lens FOV. The filter wheel can be rotated freely by the control software. Two of the filters are the red-line filter of 630.2nm and green-line filter of 557.9nm, both with a bandwidth of 1nm, and approximately 2 degrees of angle of acceptance. The fisheye lens has 150 degrees of acceptance, and within SCANDI, a series of lenses (figure 3.3) collimate light rays to 2 degrees. Hence the geometry is condensed to within the filter's angle acceptance of 2 degrees.

Note that the filters are centred at 0.2nm higher wavelengths than the wavelength of

interest. This ensures that there is no significant reduction of transmission towards the edge of the filter pass-bands. This is because the centre wavelength of the filter is lower for off-axis rays. This technique ensures that all rays within the light cone pass through the filter.

The third filter is a 632.8nm filter with bandwidth of 10nm, which is the filter for laser calibration (the LASER command), and is vital for SCANDI calibration. For a single viewing FPI, the LASER is essential in determining the instrumentation function in temperature calculation, which is also the case for SCANDI. For the all-sky imaging system, the LASER has the additional function of constructing a 'phase map', which accounts for instrumentation influences including the etalon spacing and the axial angle of individual pixels (section 3.3).

In addition to the LASER calibration, further calibration is also available from the 630.4nm line of a neon lamp (the CAL command), which is an electrode-less lamp that is excited by a radio frequency oscillator. The neon CAL lamp is especially important for a single view FPI, since it can account for any drifts in interferograms caused by instrumentation effects like room temperature variations. For SCANDI, the instrumentation drift is accounted for automatically by the phase-mapping process. Nevertheless, the CAL is still useful for other calibration uses.

Both the light from LASER and CAL are introduced through a 75 cm square fibre-optic panel, which is then directed to the system by a 45^o mirror in a position prior to the filter wheel. It can be rotated into the optical path when calibrations are performed. The position of SCANDI's filter wheel is also different from those of the FPIs, which is located after the etalon and before the imager. Four calibration images are on average taken for every 24-hour observing period, as of season 2011.

3.2.1.3 Beam expander

The emerging beam is then expanded to 150mm diameter to match the aperture of the etalon. The corresponding line cone angle is about 3^o, and its entire field of view is mapped onto five orders of interference of the etalon.

3.2.1.4 Capacitance Stabilised Etalon (CSE)

The CSE is designed and developed at APL, and is well documented in chapter 11 of McWhirter [1993]. There are in general two ways in which a FPI can change the angular

diameter of the interference pattern (i.e. to be scanned): pressure scanning, which changes the refractive index within the etalon cavity, or by changing the physical distance between the etalon plates/mirrors. Both methods vary the optical path difference in terms of wavelength. The second method is used in the CSE. Any instrumentation influences are required to be accounted for with the phase-mapping process (section 3.4).

The CSE is similar to a fixed-gap etalon, but three conventional fixed spacers separating the two parallel mirrors are replaced by piezo-electric transducers (PZT). The PZT are made from a material which changes its physical size with an applied voltage. However, difficulties are involved in monitoring the size change because the extension per volt is not consistent enough for the purpose here.

Therefore, the cavity length needs to be controlled accurately by monitoring each spacer with a capacitance sensor. The sensor is a parallel plate capacitor with one plate attached to each etalon surface. Variations in the etalon gap produce a change in the value of capacitance, which is then measured and used to adjust the piezo-electric voltage. The three spacers are then controlled through a three channel loop which incorporates the sensitive capacitance bridge circuits, and high voltage amplifiers. This is the Capacitance Stabilised Etalon. Hence the spacer length can be controlled accurately via a computer.

The etalon is also contained in a sealed, temperature controlled housing, mounted on a three-point support which allows tilting for centring the fringe pattern in the final image.

3.2.1.5 Beam reducing telescope and EMCCD imager

The final optics consists of a Galilean beam reducer and a (f/1.8) commercial camera lens. It focuses the final image onto the Electron Multiplying CCD camera (Andor DV887).

The EMCCD amplifies signal using high voltage to encourage impact ionisation. Electron multiplication occurs due to the high voltage applied. As electrons are moved from one gate to another, the high voltage causes impact ionisation where electrons hit the lattice at high energies. The impact ionisation happens in an additional row of serial readout elements compared to other CCDs. A statistical gain of 1.01 per elements can be achieved, with a final gain of $(1.01)^n$ where n is the number of elements. The gain is equivalent to a few 100 in a typical readout row.

The signal to noise (SNR) is significantly improved in an EMCCD, and it is determined by

$$SNR = \frac{P.QE}{\sqrt{P.QE + N_{Dark}^2 + N_{Readout}^2}} \tag{3.2}$$

where P.QE is the signal intensity, which is equivalent to the number of photons (P) times Quantum Efficiency (QE). The denominator of equation 3.2 represents the noise. This includes the shot noise ($\sqrt{P.QE}$), which is noise related to the peak of the signal; and the dark signal noise (N_{Dark}) and readout amplifier noise ($N_{Readout}$), which both contribute to the 'background' noise. The EMCCD is able to achieve good signal to noise ratio because it has a gain mechanism that amplifies the signal before it enters the readout amplifier i.e. $N_{Readout}$ is separated from the gain and signal intensity P.QE, which improves the signal-to-noise ratio significantly.

3.2.1.6 Optical alignment

The optics may be considered as the superposition of two separate systems: an all-sky imager and a Fabry-Perot interferometer, and each has to be aligned and focused independently to produce the final image.

The Fabry-Perot interferometer is focused using the LASER command. Laser light is injected through the appropriate laser filter, and through the etalon. This produces stable Fabry-Perot fringes, which are viewed by the EMCCD camera. First, the final camera lens is adjusted to focus the fringes, and then the fringe pattern is centred by adjusting the three support points of the etalon.

The all-sky imager is focused with broad-band sky light i.e. the empty position in the filter wheel is rotated into place, with the LASER and CAL turned off and out of the light path. In this configuration the instrument functions solely as an imager and the sky image can be centred and focused. The centring is performed by replacing the field lens at the primary image plane of the fisheye lens (figure 3.3) with a central pinhole aperture. The lateral position of the lens is then adjusted to centre the pinhole in the camera image.

3.3 SCANDI background theory and application

This section gives a summary on the essential Fabry-Perot interferometer theories, as well as the software responsible for the all-sky imaging function of SCANDI. In addition to McWhirter [1993], Aruliah et al. [2010], Conde and Smith [1997] is the core document for the phase mapping procedure.

3.3.1 Fabry-Perot Etalon

The Fabry-Perot etalon needs to be very sensitive in order to observe spectral Doppler shifts and temperature broadenings in the thermosphere. For example, a wind speed of 200 m/s detected from the 630nm red-line emission would correspond to a Doppler wavelength shift of only 4.2x10^{-13}m. For the FPI, each bin of a 256 bin readout for a 18.5mm etalon gap corresponds to 4.19x10^{-14}m per bin. Hence, the Fabry-Perot etalon is well suited for measurements in E and F-regions.

An etalon is a cavity formed by two perfectly flat and parallel mirrors with ∼90% reflectivity. As light of wavelength, λ, enters an ideal etalon of mirror separation, d, at near normal angle, θ; the optical path difference, δ, between successive reflections can be described by

$$\delta = 2dcos\theta \tag{3.3}$$

and their phase differences, $\Delta\phi$, is

$$\Delta\phi = \frac{2\pi\delta}{\lambda} = \frac{4\pi\mu dcos\theta}{\lambda_0} \tag{3.4}$$

where μ is the refractive index, and λ_0 the vacuum wavelength. The multiple reflection in an etalon cavity is demonstrated in figure 3.5.

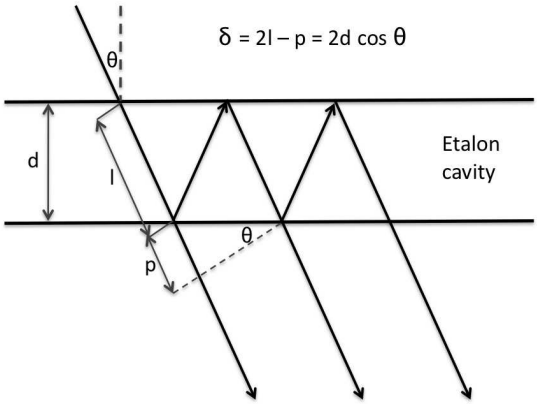


Figure 3.5: Multiple reflection in an etalon cavity. Courtesy McWhirter [1993].

The transmittance, T, of a perfect etalon of reflectance, R, is represented by the Airy function, A(n):

$$A(n) = \frac{\tau}{1 + \frac{4R}{(1-R)^2} sin^2\,(n\pi)} \tag{3.5}$$

where $\tau = T^2/(1-R)^2$, and n is the order of interference. It can be seen that λ and cos θ are interdependent, and the angular diameter of a given fringe is dependent upon wavelength. Hence the etalon can be used to measure small changes in wavelength by measuring the fringe diameter.

The factor

$$F = \frac{4R}{(1-R)^2} \tag{3.6}$$

in equation 3.5 is the coefficient of finesse, F, from which it can be seen that the higher the reflectance, the sharper the transmitted fringes. Hence, the Airy function can be simplified as

$$A(n) = \frac{1}{1 + Fsin^2\left(\frac{\Delta\phi}{2}\right)}. \tag{3.7}$$

Maximum transmission occurs at integer values of n, given by

$$n = \frac{2\mu dcos\theta}{\lambda_0}. \tag{3.8}$$

If θ, d and μ remain constant, the change in wavelength corresponding to an increase in order from n to n+1 is the free spectral range (FSR or $\Delta\lambda_{FSR}$). The FSR is a very important parameter for the etalon, since it represents the range of wavelengths which can be displayed without overlap.

Equation 3.8 can be re-expressed as 2dcosθ=nλ, which is the wavelength for a maximum transmission at order n. There will also be another maximum transmission at $(\lambda + \Delta\lambda)$ at order (n-1). This is described by 2dcos θ=(n-1)$(\lambda + \Delta\lambda)$.

Hence,

$$\frac{2dcos\theta}{\lambda} = \frac{2dcos\theta}{(\lambda + \Delta\lambda)} + 1 \tag{3.9}$$

and can be rewritten as

$$\frac{\Delta\lambda}{\lambda\,(\lambda + \Delta\lambda)} = \frac{1}{2dcos\theta}. \tag{3.10}$$

As $\Delta\lambda \ll \lambda$ and for near axis rays cos $\theta = 1$, the equation becomes:

$$\Delta\lambda_{FSR} = \frac{\lambda^2}{2d}. \tag{3.11}$$

It should be noted that the coefficient of finesse, F, described in equation 3.6 should not be confused with the reflective finesse, N_r, which is the reciprocal of the full width half maximum (FWHM) of the etalon transmission profile, normalised by the FSR.

By inspection of equation 3.7, the FWHMs of the Airy function occur at $Fsin^2\frac{\Delta\phi}{2} = 1$. With the approximation that $sin^2\frac{\Delta\phi}{2} \approx (\frac{\Delta\phi}{2})^2$ since F is usually sufficiently large; the fact that half maxima occur at +/- $\Delta\phi$, and the full width at half maximum is therefore $2\Delta\phi$; and also that a FSR corresponds to a phase change of 2π; the reflective finesse, together with the coefficient of finesse (equation 3.6), can be expressed as

$$N_r = \frac{2\pi}{2\Delta\phi} = \pi\frac{\sqrt{R}}{(1-R)}. \tag{3.12}$$

This is an important parameter because it defines the theoretical resolving power of the etalon.

There are two basic methods in which the Fabry-Perot etalon can be used to measure a shift in wavelength, which are scanning and imaging. Both methods make use of the interdependence of λ, μ, d and cos θ in equations 3.5 and 3.8.

The scanning method basically involves measurement of the intensity of light at a particular value of cos θ, usually on the optical axis of the instrument with cos θ=1. d or μ are then varied and stepped through the transmission profile to give a full Fabry-Perot fringe. The capacitance stabilised etalon in SCANDI is able to maintain the stability and parallelism of the gap while the distance between the parallel mirrors, d, is varied.

The imaging method uses the cos θ dependence, and the angular diameter changes of the fringes are measured. A single viewing FPI uses the imaging method only, while SCANDI relies on both the scanning and imaging methods.

For the imaging method, the angle subtended by a given fringe, θ, needs to be determined, and a converging lens of known focal length, f, is used to form an image of the rings. Equation 3.8 can then be re-expressed in terms of f and the radius of the fringe rings, r, as

$$\lambda = 2d\left(1 - \frac{r^2}{2f^2}\right) \tag{3.13}$$

with the approximations that $cos\ \theta = 1 - \frac{\theta^2}{2!} + \frac{\theta^2}{4!} - ...$, and since θ is small, the high orders can be neglected and $\theta \approx tan\theta$; as well as applying the simple lens theory that a circular object at infinity sub-tending a half angle of θ forms an image with radius, r, and is described by $tan\ \theta = \frac{r}{f}$.

Equation 3.13 shows that $n\lambda \propto r^2$ i.e. the adjacent maxima of a fringe, formed by consecutive integer values of n, form a series of concentric fringes. These rings have equal separation in r^2 space.

A single-view FPI, which applies the imaging method only, is able to measure an entire transmission profile at each of its viewing directions. Each transmission profile usually contains 2-3 orders of the circular fringes, and a bisector is then taken on the circular rings. The final fringe is produced by summing the entire circular fringe via the bisector, forming a 2-dimensional transmission function. The fringe data are saved in the linear r^2 space so there are equal spacing between the maximum transmissions. Usually 1-2 orders are included in the final output. Hence the final result is an averaged-profile over the 1-degree FOV of the FPI at each of its viewing directions. For SCANDI, both the scanning and imaging methods are needed to acquire the full spatial information in the 150 degrees FOV.

3.4 Phase compensation in a scanned and imaging etalon

Details about the phase compensating process can be found in Conde and Smith [1997], and the process is summarised here. This section is an important part of the thesis because it describes the major physical and theoretical differences between a narrow-FOV FPI and SCANDI.

The etalon is able to measure wavelength changes through the interdependence of λ, d, μ and cos θ. For SCANDI, which combines the scanning and imaging methods, the scanning is performed by varying the etalon spacing, d, over n discrete steps while μ remains fixed; and the all-sky imaging is performed by mapping the transmission fringes onto the final (512x512) EMCCD chip. Unlike the FPI, the photons detected at different pixel positions (x,y) represent measurements at significantly different axial angles, θ, because of the 150^0 FOV.

In other words, there are two contributions to the instrument transmission function, which are the shift in wavelength $\delta_d\lambda$ associated with each successive value of d during a scan, relative to that at the start of the scan; and the shift in wavelength $\delta_\theta\lambda$ associated with the incident angle of the measurement, relative to the transmission at normal incidence.

Hence, for each photon event, the total instrumentation spectral wavelength shift, relative to the beginning of a scan, and an on-axis pixel is

$$\delta_\star\lambda = \delta_d\lambda + \delta_\theta\lambda \tag{3.14}$$

Since the Airy function is periodic in wavelength with period $\Delta\lambda$, the difference between the spectral responses of two pixels, during two different photon detections can be determined completely by $\delta_\star\lambda/\Delta\lambda$. It represents the change in the order of interference for a photon arriving at the detector at any time and position, relative to a photon arriving at the detector on axis and at the beginning of a scan.

Therefore in practice, the spectra are accumulated as functions of the equivalent fractional order changes, i.e. frac($\delta_\star\lambda/\Delta\lambda$), denoted here as $\phi_\star$. $\phi_\star$ is the phase shift of the transmission function (equation 3.7), A(n), due to the characteristic of the instrument, and needs to be compensated. Hence, phase maps need to be constructed to account for the instrumentation phase shift for each pixel position (x, y) and at each etalon step, d. The phase compensation is performed in real time.

During spectral acquisition, $\delta_\star\lambda/\Delta\lambda$ is maintained at values smaller than 1 by controlling the range of d values that can be scanned. As d is varied, if a transmission image on the EMCCD is looked at in real time, as shown in figure 3.6 ; the maxima/rings of the transmission will contract with the phase change caused by the etalon stepping. Eventually, the whole FOV is scanned by a maxima (i.e. points of constructive interferences) at some point of the etalon stepping. Hence, each pixel is able to accumulate a complete interferogram.

The photons arriving at each pixel at all points of a scan can be summed because they will be in phase with each other after the phase compensation. If no spatial information is required, all pixels can be combined into a single interferogram. As of the end of observing season in March 2011, SCANDI's detection is summed into 25 subset of pixels i.e. zones, as demonstrated in figure 3.6.

The phase map is obtained experimentally, and all geometric instrumentation information about the fringes is contained in the phase map, and hence no explicit corrections for poor centring or non-circular fringes are required. The 632.8nm He-Ne LASER is used because it is a good calibration source that is able to produce a spectrally narrow and stable emission. With the phase map, the points at which an interferogram maximises for each pixel, and at each position of a scan can be determined. The detailed process of obtaining a phase map from a laser source can be found in section 3B of Conde and Smith [1997].

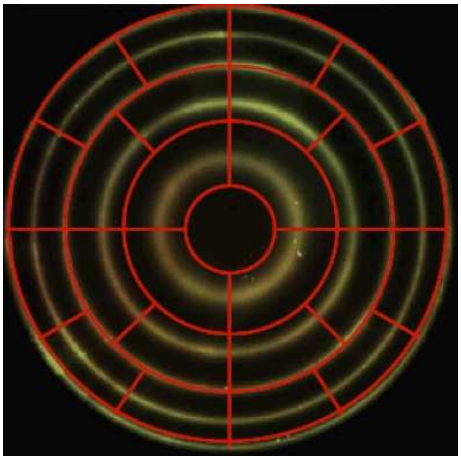


Figure 3.6: Courtesy McWhirter, I.. An imitation snapshot of an EMCCD image with the etalon rings, where the grid lines show the CCD locations of the 25 zones. Information from the sky is collected at the rings, where constructive interferences occur. As SCANDI scans, the rings expand outwards covering a full FSR, hence data from all points in the sky image can be collected. Note that this is not an actual EMCCD image, but an all-sky colour image modulated by the etalon ring transfer function, which is used here as an illustration.

The phase map is relatively stable and can account for instrumentation effects such as physical shift of optical components, degrading filter, etalon stepping divergent and changes in instrumentation temperatures.

The mechanical scanning process of SCANDI can be summarised as follow:

- SCANDI had a 10 bits etalon control system, i.e. it has a resolution of 512+512=1024 steps. So theoretically, any number of FSRs within the 1024 steps can be set. It is decided that 1 FSR should be 600 etalon steps. From 2008 onwards, the system was changed to 8 bits because it is faster.

- Also after further testing, the best results with minimal dead time and the highest resolution, can be achieved by allowing 120 steps for each complete spectral profile/ interferogram.

- Each step, associated with a particular 'section' or phase of an interferogram, takes 0.5s to expose, and 0.24s to process. The processing time includes applying the phase map to co-add equivalent pixels for each zone, as well as the time for the etalon to move for 1 step. The information within a zone therefore represents the average spectral condition of its observing region. Higher spatial resolution can achieved with smaller zone division, provided signal-to-noise is acceptable.

- For each complete cycle of 120 steps, a complete spectral profile/interferogram image is produced at each pixel, and hence at each of the 25 zones. An example of the SCANDI control screen with the 25 interferograms after a complete cycle is shown in figure 3.7.

- This is repeated for 5 cycles i.e. a total of 120x5=600 steps, and the 5 cycles are then averaged to form one final spectral profile for each zone. The total process therefore takes 7 minutes and 24 seconds. If one long cycle of 600 steps is used per spectral profile, it would give a higher possibility of intensity change from the light source over a cycle. The intensity change could induce a false peak and Doppler shift. Hence, averaging over 5 shorter cycles at 120 steps is chosen for SCANDI.

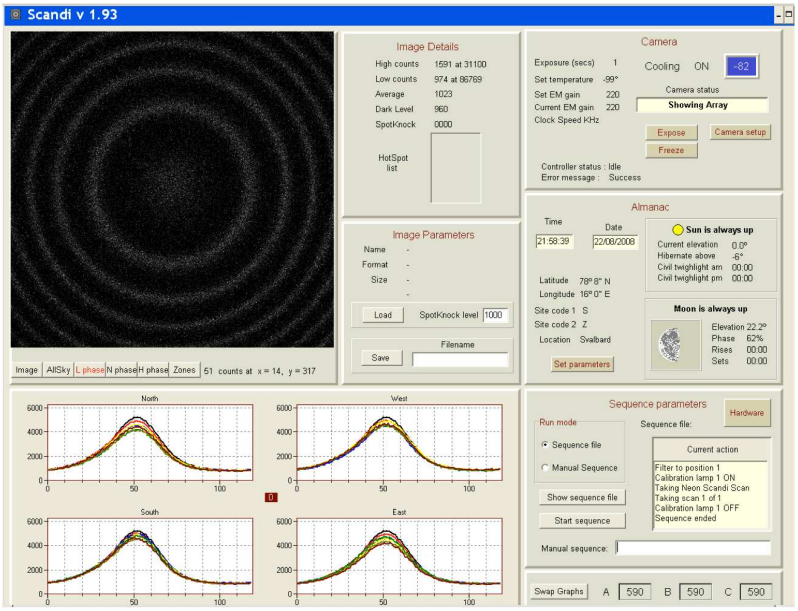


Figure 3.7: Courtesy McWhirter, I.. A SCANDI control screen that shows the 25 interferograms for each zone after one complete scan in the bottom left. Note that the 25 interferograms from each zone in each observing quadrant (NSEW) are over-plotted on top of each other within the 4 plotting windows. The slight variation in spectral shape of the interferograms are expected because they are actual measurements of the sky.

3.5 Data processing

3.5.1 Background

Three neutral thermosphere parameters can be retrieved from the measured spectra, which are the temperatures determined from the full-width-half-maximum (FWHM) of the emis-

sion line profile; the neutral wind velocity from the Doppler shifts of the spectral peak, as well as the 630nm/557.7nm emission intensity from the spectral peak height. This section gives an introduction to the basic theories and procedures involved in the data retrieval process.

At around 240km altitudes, the high percentage of oxygen and the thin atmosphere allow time for oxygen to emit red-line radiation, but collisions become more frequent towards the lower atmosphere, so that red emissions cease. Hence only green-line emission is seen at ∼110km altitude.

An emission line profile is broadened as a result of the Doppler effect, where the apparent emission wavelength is shifted if there is a component of velocity along the line-of-sight (LOS) of the observer. When the red or green-line emissions are detected by SCANDI/FPI, the line profiles are broadened by the random thermal motions of the emitting atoms, representing the temperature. It is assumed that the emitting species i.e. atomic oxygen is in thermal equilibrium with its surroundings (section 2.1.1). The overall LOS velocity of the wind is represented by the peak Doppler-shifted position of the broadened line profile.

A Doppler shift in wavelength emitted by each atom is given by

$$-\frac{\Delta\lambda}{\lambda_0} = \frac{\Delta\nu}{\nu_0} = \frac{v_x}{c} \tag{3.15}$$

where v_x is the LOS velocity, c the speed of light, λ the wavelength, and ν is the frequency. The random motion of the atoms in the gas can be expressed as follows with simple kinetic theory:

$$\frac{1}{2}m\overline{v^2} = \frac{3}{2}kT \tag{3.16}$$

where m is the mass of an atom, k the Boltzmann constant, T is the absolute temperature and $\overline{v^2}$ is the mean square velocity of the atoms in the gas. Further details can be found in e.g. Griffin [2000], Thorne [1988], the velocities of the atoms can be described by the Maxwell-Boltzmann distribution (section 2.1.1):

$$f(v) = 4\pi \left(\frac{m}{2\pi kT}\right)^{\frac{3}{2}} v_x^2 \; exp\left(\frac{mv_x^2}{2kT}\right). \tag{3.17}$$

The LOS velocity, v_x, can be substituted with the Doppler equation (equation 3.15) and expressed in terms of frequency and wavelength. The Maxwell-Boltzmann distribution is a

safe assumption under most circumstances because the lifetime of an excited thermospheric atomic oxygen atom is about 110 s and 1 s for red and green-line emissions respectively. The ion-neutral collision frequency is about 1 s^{-1} and 500 s^{-1} respectively for the F and E-regions. Therefore the lifetimes of both emissions allow the atomic oxygen particles to reach thermal equilibrium with their surroundings.

The Maxwell-Boltzmann distribution can also be written in terms of the fraction of number of atoms within a certain velocity range. Since the intensity, I_v, is proportional to the number of atoms, the emission spectra can be described in Maxwell Boltzmann terms as

$$I_\nu = I_0 \; exp\left[-\left(\frac{c}{\nu_0}\right)^2 \cdot \left(\frac{m}{2kT}\right) \cdot (\nu_0 - \nu)^2\right] \tag{3.18}$$

where I_0 is the peak intensity. The FWHM occurs at $I_\nu = I_0/2$, therefore it leads to the FWHM expression, $\delta\nu_D$:

$$\delta\nu_D = \frac{2\nu_0}{c}\sqrt{\frac{2RT\ ln2}{M}} \tag{3.19}$$

where R is the ideal gas constant and M is the mass number for the emitting species. Its dimensionless form is

$$\frac{\delta\nu_D}{\nu_0} = \frac{\delta\lambda_D}{\lambda_0} = 7.16 \times 10^{-7}\sqrt{\frac{T}{M}}. \tag{3.20}$$

Hence the temperature can be deduced from the emission profile detected from the FWHM of the profile, wind from the peak position of the profile, and intensity from the peak height.

3.5.2 Data reduction

The SKY emission spectra, as well as the instrumentation calibration function produced from the LASER (a He-Ne laser at 632.8nm) are stored in intensity vs. r^2 format, where the intensity spectra are linearly spaced. r is the radius of a point of measurement from the centre of the optical axis (section 3.3). The instrumentation function is essential for the deconvolution of the instrument's measured SKY function to produce a 'pure' emission spectra for the determination of neutral parameters, especially for the temperature.

For SCANDI, a LASER instrumentation function is measured for each zone, and the function is updated 3-4 times for a 24-hour night. The LASER peak position is also used to determine how much the SKY spectral peak has shifted, in order to deduce the Doppler

shift. The zero-Doppler shift position then needs to be found to give the final LOS wind velocity.

Unlike the FPI, the LASER peak position is essential in the Doppler shift determination, because of the spectral phase shifting involved in the phase mapping process (section 3.4) in SCANDI. For the FPI, phase mapping is not part of the data reduction process. The FPI LASER instrumentation function is only applied in the deconvolution process and the CAL calibration function is used to account for any instrumentation peak drift. The FPI CAL comes from a different calibration source than the LASER (section 3.2.1.2). A CAL image is taken with every directional cycle in the FPI.

The laser and calibration lamps together provide continuous monitoring of all of the factors that influence the stability and performance of the instrument.

3.5.3 Setting up SCANDI for different wavelength measurements

The LASER, CAL, and red and green SKY functions have respective wavelengths of 632.8nm, 630.2nm, 630.0nm and 557.7nm. It is necessary to scale the individual functions for the deconvolution and peak position determination processes. The usual assumption is that the difference between the LASER and red-line wavelength is negligible and no shifting is required. However, for the green 557.7nm SKY measurements, scaling is necessary because the instruments do not have a green LASER. The scaled 632.8nm LASER, which is also used for the 630.0nm SKY measurement, is a suitable substitute as instrumentation function for the 557.7nm SKY measurements.

The LASER function is assumed to be a perfect Gaussian with the form:

$$f(x) = exp - \left(\frac{x^2}{2\sigma\ a}\right) \tag{3.21}$$

where x is the horizontal extent of the LASER Gaussian in r^2 space, and σ, the standard deviation of the Gaussian relating to the width of the function. The extra factor, a, is the scaling function, determined simply by

$$a = \frac{\lambda_{laser}}{\lambda_{green}} = \frac{632.8}{557.7}. \tag{3.22}$$

Hence, at the beginning of the data retrieval process for the green-line data, the laser profile is first normalised and the factor $\frac{x^2}{2\sigma}$ in the laser Gaussian function is divided by the factor, a. The width σ is scaled with the wavelength ratio.

It is also necessary to change the number of steps in the SCANDI capacitance stabilised etalon (sections 3.2.1.4) stepping control software (section 3.3), in order to construct the Fabry-Perot interferograms correctly during the scanning process. From the bullet points at the end of section 3.3, the etalon is set to require 600 steps for each complete scan. This process scans through 5 free-spectral-range (FSR) of the LASER 632.8nm emission i.e. each LASER spectral profile is deduced by 120 etalon steps/data bins. The phase map (section 3.4) is then constructed with the LASER data. Hence it is necessary to scale the measured red and green SKY wavelength interferograms to match with the LASER phase map. The control software has an in-built function that determines the number of etalon steps that are associated with each SKY wavelength's spectral profile. This process is described as follow:

As light enters the etalon, it is reflected by the pair of etalon mirrors and the wavelength change is $1/2\lambda$ if the etalon gap changes, represented by the Fabry-Perot relationship for optical path difference in terms of wavelengths and order of interference:

$$n\lambda = 2\mu d cos\theta. \tag{3.23}$$

This is simply a rearrangement from equation 3.8, where the variables of the stepping process are the etalon gap, d, and the transmission wavelength, λ. Therefore, the number of steps per FSR is determined by the ratio:

$$\frac{\lambda_{sky}}{\lambda_{laser}} = \frac{x_{sky}}{120} \tag{3.24}$$

where λ_{sky} can either be the red or green-line SKY wavelength of 630.0nm and 557.7nm respectively, λ_{laser} is 632.8nm and 120 is the number of etalon steps for one FSR of the LASER. Hence the number of etalon steps for one FSR of the red and green-line are 119.5 and 105.8 respectively. These are rounded up to 119 and 105 in the control software. The numbers are equivalent to 4 and 5 complete measured spectral profiles for the total 600 steps in the scanning process.

It should be noted that since the etalon stepping is a physical distance related to the path difference of the incoming rays of an emission (equation 3.23), it should not be confused with the FSR expressed in terms of wavelength shift i.e. $\Delta\lambda_{FSR} = \frac{\lambda^2}{2d}$ (equation 3.11). Therefore the scaling is linear with λ and not λ^2.

Finally, for the determination of LOS wind, which is represented by the shift in peak position of the spectral profile; it is necessary to determine the velocity equivalent of the

shift in bin number. The two equations involved are the FSR in terms of wavelength shift, $\Delta\lambda_{FSR} = \frac{\lambda^2}{2d}$, and the Doppler shift equation 3.15, $\frac{\Delta\lambda}{\lambda_0} = \frac{v_x}{c}$. Together, they give the expression:

$$v_{bin} = \frac{1}{n_{bin}} \cdot \frac{\lambda_{sky} c}{2d} \tag{3.25}$$

where v_{bin} is the velocity equivalent to a single bin peak shift, and n_{bin} is the number of data bins per FSR or complete spectral profile. This is the same as the number of etalon steps for SCANDI, which are 119 and 105 for red and green-line respectively. The velocity per bin for red-line is 42.9 m/s and for green is 43.0 m/s. The whole spectral profile is fitted to determine the peak position and the resolution is not limited by the bin increments.

3.5.4 Summary of data retrieval software and error estimation

Algorithms using the IDL development language based on M. Conde (private communication) have been developed for the spectral profile fitting and atmospheric parameter determination. Technical details can be found in Conde and Smith [1997] and this section provides a basic summary to the routines.

A measured SKY spectrum and a LASER instrumentation function are given by each SCANDI zone. Therefore, for each SCANDI zone, the spectral width, peak wavelength, and source intensity along with an estimate of the background intensity of the SKY spectrum are determined. The method used is a nonlinear least-square curve-fitting algorithm. Figure 3.8 shows a flow chart summarising the data analysis algorithms.

For each observing zone, experimentally-determined LASER instrument functions are convolved with trial Gaussian emission functions, this gives the estimated FPI spectral model. The goodness of fit of the transform of each trial function, to the transform of the measured SKY data is estimated by calculating the χ^2 residuals (equation 3.26).

$$\chi^2 = \sum_{n=0}^{N-1} \frac{(y_n - s_n)^2}{\sigma_n^2}. \tag{3.26}$$

The guessed parameters that gives a χ^2 value within an acceptable statistical limit is taken as the best fit and is used to compute the parameters of interest for each SCANDI zone. In other words, the best choices of guessed parameters give a minimum value of χ^2, with an expected value of N minus the number of free parameters in the model. Hence, the optimum results are provided when:

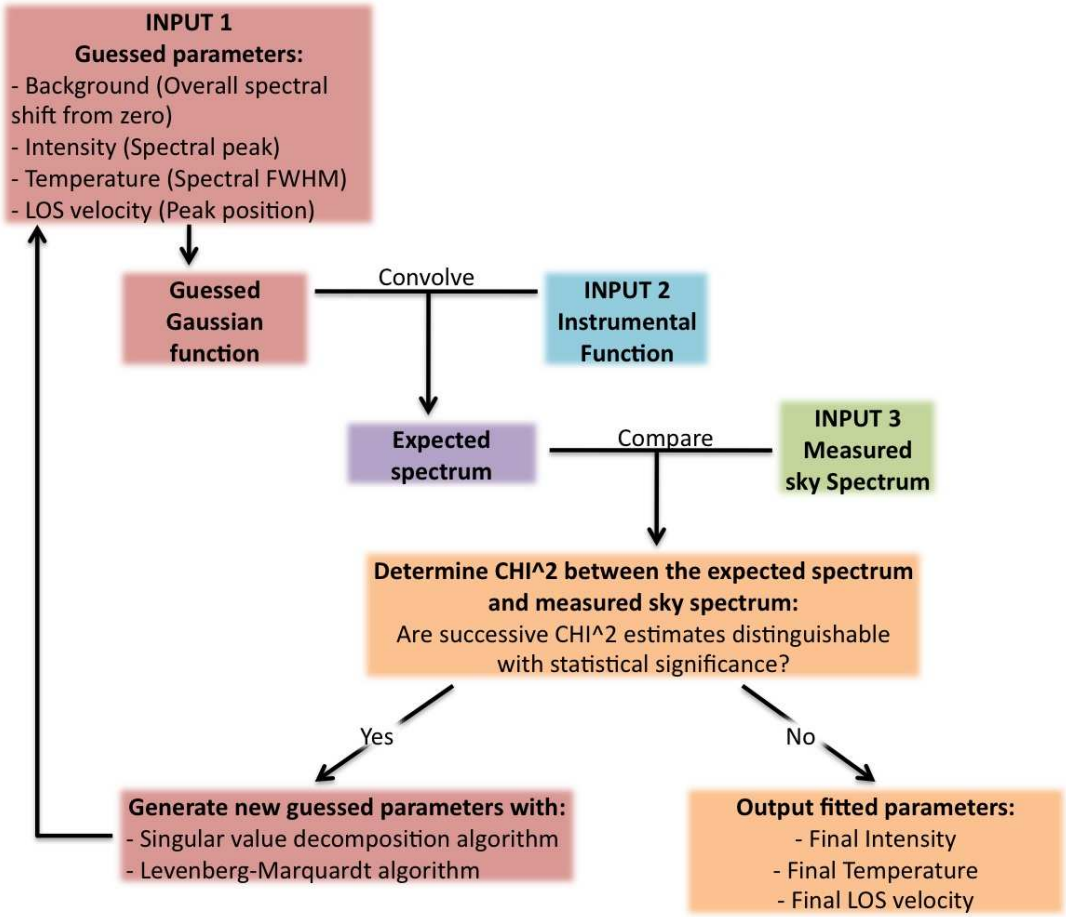


Figure 3.8: This flow chart summarises the algorithms used in the data reduction process for SCANDI. This process is applied to each SCANDI data zone.

$$\frac{\delta \chi^2}{\delta p_j} = 0. \tag{3.27}$$

p_j is the set of parameters that needs to be determined. In our case, the parameters are the background, intensity, temperature and the line of sight velocity, which is represented by the characteristics of the spectra (section 3.5.1). However, for the purpose of the FPI spectral analysis and its spectral model, it is not possible to solve equation 3.27 for all the parameters. Mark Conde constructed an algorithm which overcomes this problem by:

- solving analytically for parameters appearing linearly in the FPI model function, s_n i.e. the background and spectral intensities, with the Singular Value Decomposition (SDV) algorithm and
- solving separately for non-linear parameters i.e. the temperature, with the Levenberg-Marquardt (LM) algorithm.

These two parts are solved iteratively as described by the flow chart (figure 3.8).

It is also therefore not possible to directly compute independently the variance, and hence error, of each parameter with the SDV and LM methods. Instead, a standard non-linear least-squares fitting technique, e.g. Bevington [1969], Press et al. [1987], is used to compute the variances of all the parameters of interest simultaneously. The detailed mathematics of the algorithms for data analysis is beyond the scope of this thesis. The technical description can be found in Conde and Smith [1997], as well as 'Analysis of Fabry-Perot spectra of lidar backscatter echoes' by M. Conde, 1999.

Finally, by combining the results from each zones, it is possible to produce maps of each parameter across the entire SCANDI field-of-view. As a quality check, the SCANDI wind, temperature and intensity deduced from the SCANDI algorithms are compared against the co-located and independent FPI wind results (section 3.8.6.8).

Please note the current technique minimises chi square. At the time this thesis was written, a threshold limit is not yet applied on the analysed data. A threshold could be implemented into the algorithms as an extra quality check in the future.

3.6 SCANDI data analysis and calibration introduction

The following sections describe the calibration processes, which include the all-sky projection method and the calibration for the major SCANDI neutral parameters: temperatures, line-of-sight winds and the emission intensity. They are based on sections 2.3 (sky mapping) and 2.4 (SCANDI data calibration) in Aruliah et al. [2010] and are updated with developments as of season 2011. They are the author's original work.

In addition, the line-of-sight wind data go through a horizontal wind fitting process, which gives 'snap-shots' of horizontal neutral wind field over the instrument FOV. The horizontal wind field fitting methods are based on Conde and Smith [1997, 1998]. Its theory and calibration processes are also described.

3.7 Instrumentation calibration - Sky mapping

3.7.1 Fish-eye lens projection onto detector

SCANDI uses a fish-eye lens in its all-sky imaging optical system to project a 150 degrees field-of-view onto its EMCCD. There are various types of manufactured fish-eye lenses available in the market and their relative performance can be found in Kumler and Bauer

[2000]. An equi-solid-angle projection fish-eye lens was chosen for SCANDI, and each incoming ray can be mapped onto an output image by the model:

$$r = 2fsin\left(\frac{\theta}{2}\right) \tag{3.28}$$

where r is the radial distance between the image points and the optical axis, f the fish-eye lens focal length, and θ the angle between the incoming ray and the optical axis (figure 3.4). Equation 3.28 is the 'perfect lens' version of the fisheye lens equation described by equation 3.1 (section 3.2.1.1), with 2f=α and 0.5=β. Since SCANDI has a vertical optical axis, θ is effectively the zenith angle.

To allow proper projection of SCANDI data onto the thermosphere, it is necessary to quantify how an all-sky image is translated through SCANDI's optics onto the detector, and to allocate the corresponding azimuth and elevation observational angles to each EMCCD pixel and SCANDI zone.

There are two possible sources for the projection calibration curve: the detected moon and star-field positions. Both can be compared against the corresponding locations from, for example, an almanac or star charting software such as Skymap (www.skymap.com). Both sources are tested and described.

The equisolid-angle model (equation 3.28) alone is not sufficient to represent the whole projection system, which is an ideal derivation from the reflection of a reflective sphere only. All real lenses would deviate from it by a few percent. Furthermore, there are other optics between the fisheye lens and the detector that need to be considered. Hence equation 3.28 is used as a guideline only, and a calibration projection curve is independently fitted in the end.

3.7.2 Centring

The optical axis does not lie exactly at the centre of the EMCCD detector. The star charting software can be used to determine the azimuths, ϕ, of each star detected on the EMCCD, at pixel position (x, y) with (0,0) at one of the corners of the chip. The centre of the optical axis/zenith on the EMCCD, (x_0, y_0), and each pixel position and azimuth from a detected star are interdependent on each other:

$$\frac{(x - x_0)}{(y - y_0)} = tan(\phi). \tag{3.29}$$

For each star on the EMCCD, ϕ is deduced from the Skymap, and the pixel position (x, y) can be noted simply by its position on the chip. The Skymap azimuth angle is relative to the zenith determined from the observer's position (not the celestial zenith). It is assumed that there are negligible differences between the Skymap's observer zenith and the optical zenith of SCANDI. Hence, after a sufficient number of star/points are collected, the zenith optical axis on the EMCCD image i.e. (x_0, y_0) can be determined. 100 star points are picked over 24 hours and the rotation of sky field allowed a wide range of azimuth to be used. Only the clearest star points that can be spotted clearly by eye on the image are used.

This centre check is performed on both of the curve fitting processes using the moon or star data. This is important because the curve fitting looks for a least square fitted relationship between θ and r. Different (x_0, y_0) therefore affects the r values. The results here could be further applied in the construction of the phase map (section 3.4), so that an appropriate centre is used. It was found that the optical axis falls on pixel position (260,250) for the SCANDI configuration in season 2011.

3.7.3 Moon

An empty location in the filter wheel enables the instrument to be operated in broadband mode. The moon and star-field, which are quasi-white light sources, can be imaged without the added modulation of the Fabry-Perot rings. The moon data are tested first. The moon appears as an elongated bright circle on the detector, shown in figure 3.9. As the night progresses, the moon rotates and appears at different pixel positions. Its azimuth and zenith angle are noted from an almanac.

Fourteen hours of static fringe images were taken from SCANDI on 24^{th} December, 2006, and 50 moon positions were chosen for the calibration. It is investigated by linearly fitting the fish-eye model (equation 3.28) to the data and the results are shown in figure 3.10. A linear fit is performed on the results as y=ax+b, where a=2f*n and n is an extra fixed reduction factor due to the other optics in SCANDI, and b is a systematic offset from the EMCCD centre.

The linear fit results give f=3 mm, with a systematic offset of 47 pixels. However, f=3 mm is more than 50% different from the value provided by the lens manufacturer, f=7.82mm. Hence the moon is unlikely to be a good source to provide an accurate projection calibration.

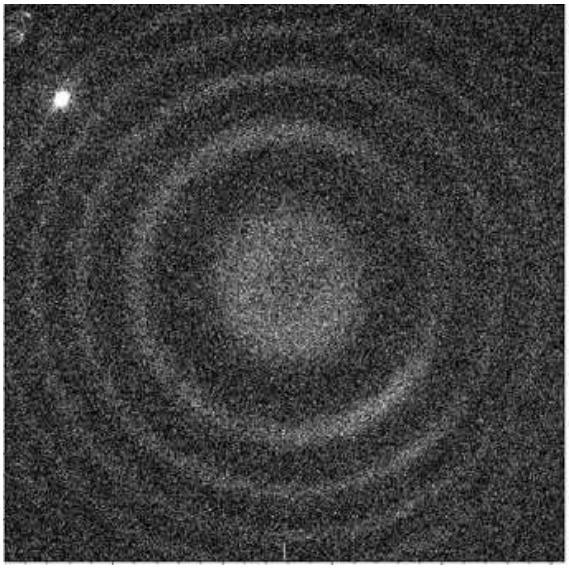


Figure 3.9: Courtesy McWhirter, I.. A EMCCD fringe image with the projected moon as a bright spot on the top left corner.

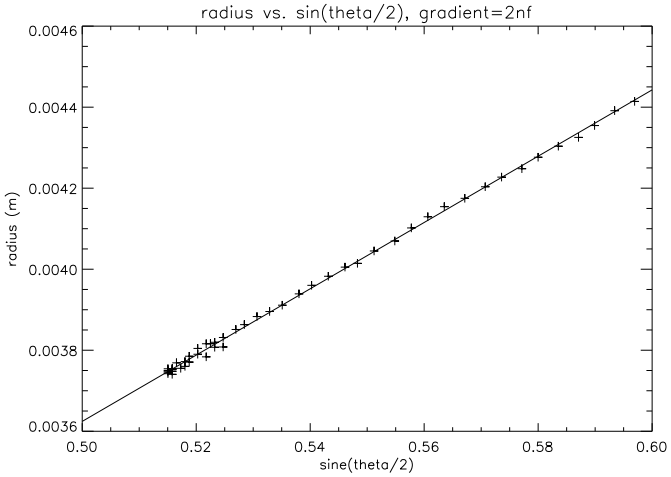


Figure 3.10: A plot of r against $\sin(\theta/2)$, where the gradient represents 2f*n (n=8/18 is an extra fixed reduction factor due to the other optics in the SCANDI all-sky system). The crosses are plots of the radial distances between the moon image points and the optical axis, against the almanac elevation angles substituted into $\sin(\theta/2)$. The straight line is the best linear fit, y=ax+b, where a=2f*n and b a systematic offset from the EMCCD centre. Note that the observed zenith=θ=90-elevation.

There are several problems associated with the moon data itself. The shape of the moon is altered into an oval by the fringes, and the bright spot on average has a 12 and 20 pixel range in the x and y directions respectively. So it is difficult to pinpoint the exact moon centre. Another difficulty is the small range of azimuths (120 degrees) and

elevations (11 degrees) available from the moon data. In addition, fitting the data against a simple theoretical equation 3.28 is not the best approach, which does not account for any instrumentation effect.

3.7.4 Star field

Star-field images are constructed for 29th February, 2008. Please note the data set here is different from the data used in Aruliah et al. [2010]. This is because SCANDI was first installed in the Auroral Station for season 2006/2007, but was moved to the Kjell Henriksen Observatory from season 2007/2008 onwards. It is expected that the transport process would shift the instrument optically by small amounts, and the star-field data set here is updated for the latest SCANDI installation.

The star-field was clearly visible from most images, allowing a high number of calibration points, with a good range of elevation and azimuth angles. 100 star locations were collected for the projection fit. Two equations were fitted for comparative purposes: a least-square fit and the generalised fish-eye lens model.

A least square fit was performed directly on a radius versus zenith plot, shown in figure 3.11. Each pixel is 1.6×10^{-4}m wide and the 512x512 EMCCD chip is 8.2 cm across. The fit is good with a χ^2 value of $\sim$0.1, and the complete fit equation is given by:

$$r(\theta) = \left(-1.36 + 383\ \theta + 29.3\ \theta^2\right) \times 10^{-5}. \tag{3.30}$$

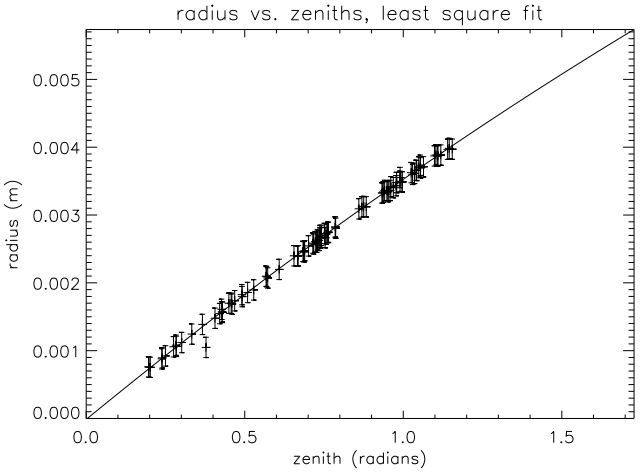


Figure 3.11: The least square fit (solid line) on the plot of radius (distance from a pixel from the optical axis on the EMCCD) vs. zenith angle (obtained from sky map). Note that the error bars are the uncertainties of the CCD pixel position of a star, assumed to be +/- 2 pixels.

A direct comparison can be made with equation 3.1, where α is the scale factor parameter, and β the radial mapping parameter. The radial distance from the optical axis on the EMCCD is matched against the radius calculated by the model; and best fit values of α and β are shown in table 3.1. These values are compared with values obtained by Thoby [2006] for a similar Sigma 8mm fisheye lens, and the ideal parameters in the simple fisheye equation (equation 3.28). The differences between the best fit and the Thoby [2006] parameters are likely to be the results of the other optics in the SCANDI all-sky imaging system. As can be seen from the table, the differences are very small. Eventually the direct least squares fit was chosen as a more precise fit to model the fisheye lens projection in SCANDI, and the azimuth and elevation angles associated with each pixel and SCANDI zone were found for future references.

Table 3.1: A table comparing the best fit and expected results. The expected values come from Thoby [2006], where an experiment was performed on the same Sigma 8mm fisheye lens as SCANDI. This can be further compared with the coefficients in equation 3.28, shown in the last column of the table.

	Best fit (Eq 3.1)	Thoby [2006] parameters	Simple fisheye lens parameters (Eq 3.28)
α	2.28	1.88	2
β	0.46	0.54	0.5

3.7.5 Look-direction projection onto detector

For the purpose of inter-comparisons with other co-located instruments, SCANDI's 25 zones are projected onto a geographic map, as shown in figure 3.12.

An altitude of 240km was chosen for the projection because it is the approximate altitude of the 630nm emission peak. The emission roughly follows the expansion and contraction of the thermosphere over a solar cycle and likely to peak at constant pressure level, instead of a fixed height. Individual geomagnetic events can also cause short term changes in the emission altitude of around 50 km in time scale of 10s of minutes. The altitude variations are studied in detail in section 4.2.3. A fixed 240km is chosen for the map projections in this section.

The equivalent look directions at 240km are calculated by simple geometry with equation 3.31, where R_{240} is the radial distance from the centre of SCANDI's FOV at the height

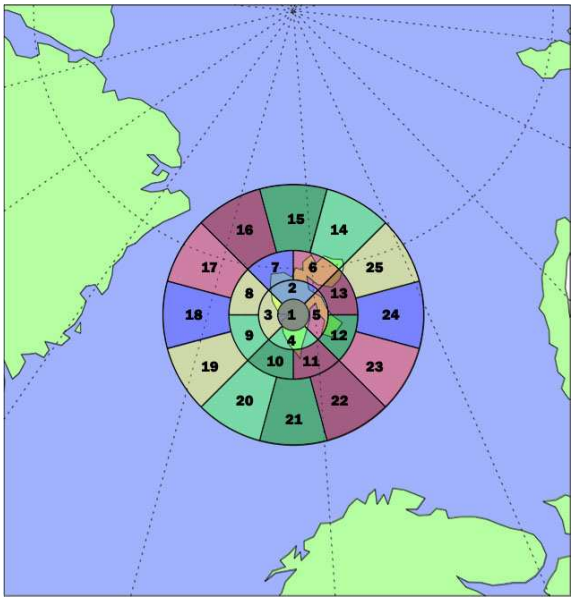


Figure 3.12: SCANDI zones projection at 240km altitude, where top is north. This configuration is equivalent to dividing each 'ring' of the zones evenly by equal radial distance on the EMCCD chip. The setting here is appropriate for winter 2006/2007. This method gives larger projected outer rings, then the inner ones. The volume of emission being measured is hence also different for each ring.

of the emitting volume, h. Azimuthal equidistant projection from the mapping function in IDL (IDL 6.3 online help, March 01, 2006) is used to provide the positions of the 240km zone-grid above Svalbard. If the curvature of the Earth is considered, the area covered by each zone and R_{240} would be different. It is applicable for the calculation of the volume of the red-line emission layer that is being measured. However, for the mapping purpose here and with the IDL map projection, equation 3.31 is sufficient to map SCANDI zones on the CCD chip to the sky.

$$R_{240} = h\ tan\theta \tag{3.31}$$

Different zone divisions are also tested. In the first winter of operation in 2006/2007, the 25 zones are divided according to approximately equal areas on the detector's EMCCD chip. After they are projected onto 240km altitude, zones from the outer SCANDI rings are elongated, with larger projected area than those of the inner rings, as demonstrated by figure 3.12. The inner and outer zenith for each ring, the associated solid angles and projected area are listed in table 3.2.

For data analysis purposes, for example, to achieve equal weighting of all zones when

Table 3.2: A table showing the inner and outer zenith angles for each SCANDI ring, and the associated solid angles and projected area. This is applicable for the first winter 2006/2007 when the SCANDI zones were divided according to equal area on the EMCCD chip.

	Season 2006/07			
	Inner zenith angle θ_1 (degrees)	Outer zenith angle θ_2 (degrees)	Solid angle (steradians)	Projected area Flat (km^2)
Centre	0	16	0.24	16 100
Inner ring 1	17	32	0.18	15 100
Middle ring 2	33	49	0.15	22 900
Outer ring 3	50	67	0.14	69 200

taking averages and for comparisons with other co-located instrumentation; the SCANDI field-of-view is divided according to equal projected area. As a consequence, the solid angle decreases as the distance from the image centre increases. The 25-zones equal projected area scheme is shown in the top row of figure 3.13. The left plot of the top row shows the size of projected areas on the sky, and the right plot shows the zone division for the EMCCD pixel position of the same scheme. The correct pixel positions on the EMCCD which give the desired zone division are carefully calculated with the fish-eye lens projection (equation 3.30).

Table 3.3 shows the zenith angles for each ring from winter 2007/2008 onwards. The derivation of the solid angle and the consequent area projected on the sky is shown in Appendix D.2. The last column in the table also shows the projected area if the curvature of the Earth is taken into account, with the approximation that the Earth is spherical, and the emission layer is simply a larger sphere surrounding Earth's surface at 240km altitude. Appendix D.1 shows the calculation for the curved projection. The differences are relatively small for the inner ring at ∼1%, to a larger difference for the outer ring at 27%. However, the effect of this on the temperature and wind data is expected to be small for the 25-zone resolution in SCANDI in season 2011. Furthermore, only the relative intensities are used in this thesis. So the flat emission layer estimation is expected to be sufficient. However, a curved projection needs to be considered for analysis that applies the higher spatial resolution.

The horizontal spatial resolution from a 25-zone grid is on average ∼200km, as indicated by the top left plot in figure 3.13. The 25-zone setting is suitable for low intensity

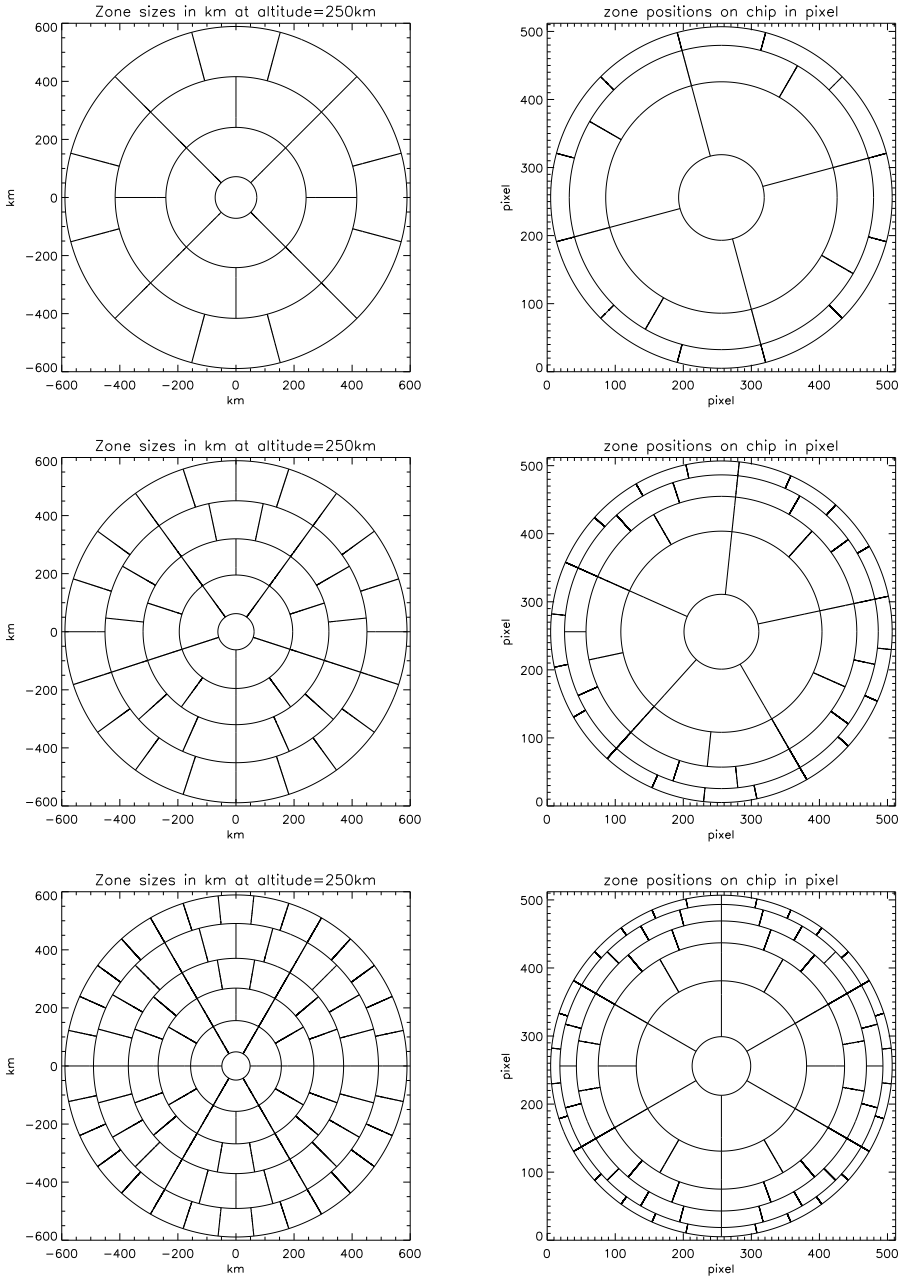


Figure 3.13: The left column shows 3 different spatial resolutions projected at 240km: 25, 51, 91 zones. The different resolutions are applicable in the associated intensity conditions, which are low, medium and high respectively. The right column is the equivalent zone grids on the EMCCD chip in pixels scale. The EMCCD chip orientation is rotated ∼30 degrees anti-clockwise with respect to north. This is because the instrument mount is aligned with the N-S geomagnetic meridian as dictated by the alignment of the observatory building, while the EMCCD zone map pattern is aligned with geographic north.

Table 3.3: A table showing the inner and outer zenith angles for each SCANDI ring, and the associated solid angles and projected area. This is applicable for the first winter 2007/2008 onwards when the SCANDI zones were divided according to equal projected area on sky.

	Season 2007/08 onwards				
	Inner zenith angle θ_1 (degrees)	Outer zenith angle θ_2 (degrees)	Solid angle (steradians)	Projected area, Flat (km^2)	Projected area, Curve (km^2)
Centre	0	16	0.24	16 100	16 100
Inner ring 1	17	44	0.38	41 700	40 200
Middle ring 2	45	59	0.16	45 100	39 600
Outer ring 3	60	67	0.065	45 500	34 600

emission for the extended period of solar minimum (cycle 23/24). However, it is possible to increase the spatial resolution for small-scale structure investigations: two new grids with 51 and 91 zones have also been constructed, with respective resolutions of ∼100km and ∼50km. The grids are shown in the middle and bottom rows of figure 3.13, and the associated zone division on the EMCCD are also presented. These grids and the associated phase map reductions (section 3.4) can be implemented when the emission intensity becomes higher for solar maximum conditions. For example, the University of Alaska SDI applies 115 zones for season 2011.

It can be seen from figure 3.13 that the associated projections on the EMCCD chip give low pixel number coverages for the narrow outer rings, and hence poorer signal to noise ratios. In theory, at lower elevations, the line-of-sight look direction passes through a greater depth of the emission layer, so more photons should be seen, but this is counteracted by increased atmospheric extinction. Also, an increase in noise is clearly seen for the outer rings for the current 25 zones setting. Therefore, it is necessary to determine an optimum zone division that balances spatial resolution and noise. However, this has minimal effect on the current 25-zone setting appropriate for the (extended) solar minimum condition over the period of work described in this thesis.

Overall, the 25-zone setting gives sufficiently large zones for a reasonable signal-to-noise ratio for the quiet solar minimum period. How the zones are set up is mostly dependent on signal-to-noise. There is the flexibility of changing the zone density to any format if the signal quality allows. The zones also do not have to be distributed evenly if only a

particular region within the field-of-view is of interest.

3.8 Data quality check

There are a number of calibration processes that are discussed in this section, which focus on the three measured parameters, emission intensity, LOS wind and temperature. The LOS wind requires particular attention, because it involves the determination of the zero Doppler shift position (baseline), as well as establishing the horizontal wind field. Please note the work described in this section is also original work by the author. The exception is section 3.8.4, where the solution is provided by Dr. Ian McWhirter.

3.8.1 Emission intensity

3.8.1.1 Atmospheric extinction and emission layer volume

The emission intensity can be deduced from the measured spectral peaks from SCANDI. At the time of this thesis, the intensity is not calibrated in Rayleighs. There are other well-calibrated emission measurements available from instruments such as the IMAGE Meridional Scanning Photometers (MSP). Nevertheless, SCANDI is still able to provide high spatial resolution of relative intensity variations over its observing map. The relative intensity measurement is also necessary to account for the intensity differences, as well as the different atmospheric extinctions, within the detected spatial volume.

The atmospheric extinction can be estimated from the distance of the emission volume from the detector, h_{obs}, which is simply the emission layer height, h_s, for the zenith zone. For the outer SCANDI observing 'rings', the distance is deduced from a spherical Earth and emission layer model (Appendix D.1, figure D.1), by $h_{obs}=(hs - h)/cos\ \theta_s$, where h is the height of the spherical cap of the whole of SCANDI FOV, and θ_s is the zenith angle of the SCANDI ring. Assuming h_s=250km, $\theta_s=67^0$ for the edge of the outer ring and h=22km deduced from equation D.6, they give h_{obs}=580km. Hence the atmospheric extinction more than doubled for the outer zones relative to the zenith measurements.

In terms of the emission volume, the FWHM of the emission layer varies and averages to about 40km. The zone volumes can be determined in a similar way from the spherical cap method described in appendix D.1. Furthermore, it is assumed that the layer is an uniform 40km-thick emission layer centred around 250km altitude. Therefore, instead of the spherical cap equation, the spherical cap volume equation is applied:

$$V_{cap} = \frac{1}{3}\pi h^2 (3R - h) \tag{3.32}$$

where R is the Earth's radius plus the height of the emission layer. The minimum volume is found at the zenith and the maximum volume is at the outer ring, which are 2x10^3 km^3 and 333x10^3 km^3 respectively for each zone. The differences in emission volume and distance from the detector are summarised in table 3.4.

Table 3.4: A table summarising the emission volume of zones at different zenith angles, assuming a 40km thick red-line emission layer. In addition the distance of the emission volume from the detector, with consideration to a curved emission layer is also shown. This is applicable to the season 2007/08 instrument setting, which applies a flat equal area projection.

	Inner zenith angle θ_1 (degrees)	Outer zenith angle θ_2 (degrees)	Distance from detector (km)	Emission Vol. (km^3)
Centre	0	16	260	1980
Inner ring 1	17	32	340	58 500
Middle ring 2	33	49	460	193 600
Outer ring 3	50	67	585	336 800

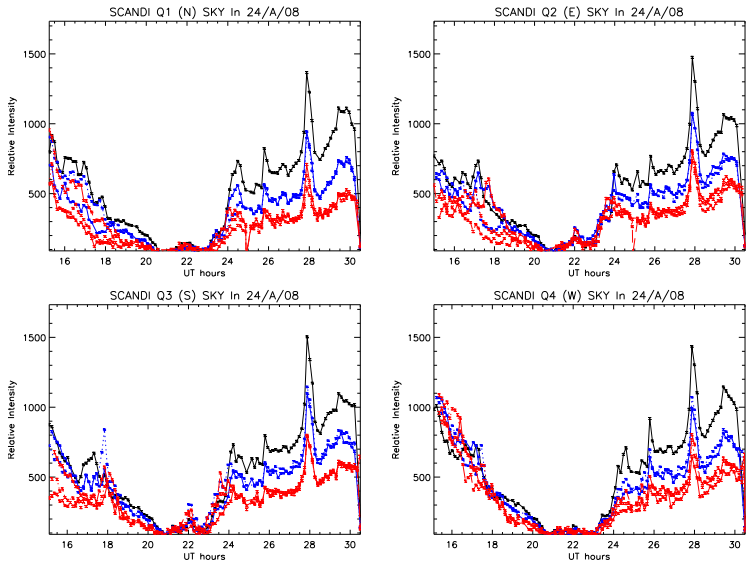


Figure 3.14: This figure shows the SCANDI-measured emission intensity, and each plot represents data from a quadrant of the SCANDI map in north (top left), south (bottom left), east (top right) and west (bottom right). An example of a SCANDI map can be found in e.g. figure 3.13. Each quadrant contains one zone in the inner ring plotted in black, two zones from the middle ring in blue, and 3 zones from the outer ring in red. The central zone is not included in this plot.

It can be seen from figure 3.14 that there is a clear distinction between intensities measured from the different SCANDI rings when no normalisation is applied to the data. It also shows that measurements from the inner ring has a higher magnitude than the outer ring's results. The results here suggest that atmospheric extinction plays a more important role than the emission volume for the determination of the true magnitude for intensity measurements.

The zenith intensity is also directly compared against the co-located IMAGE Meridional Scanning Photometer (MSP) data (not shown), which shows that SCANDI zenith intensity magnitude correlates well with the MSP results. The relative intensity measurements within a SCANDI map should be able to show features such as the location of an auroral arc, providing the correct normalisation of intensity between the different SCANDI zenith rings is implemented. In order for SCANDI to account for intensity offsets, an intensity baseline is deduced for each SCANDI ring. The baselines used are simply the average intensity over each ring for each night. It is expected some uncertainties are introduced using the intensity baselines. The variations of the baselines are therefore monitored over each observing season to account for any anomalies.

3.8.2 Temperature

When the first temperature results were deduced, there were two main temperature effects in the data that were either instrumentational or were an artifact of the data analysis program (section 3.5.4):

- the temperatures have a 20-30K offset between different SCANDI rings, and
- the zenith zone results are unusually low with an approximate 500K difference.

3.8.3 Temperature central zone effect

3.8.3.1 Diagnostic

A diagnostic process is performed with the calibration lamp (CAL) data, which show that similarly low central temperatures are present in the final reduced CAL data. The CAL data come from a 630.2nm light source (see section 3.3) and the light source imitates the red-line emission measurements, but without the temporal variations. It goes through the same optics as the measured SKY emission source and is a good indicator of instrumentation fluctuations.

The CAL results are analysed with the same data reduction process as the SKY data. Hence, seeing the same zenith zone effect for both SKY and CAL does not eliminate the possible instrumentation nor data analysis contributions to the low central zone temperature. However, it is unlikely that any instrumentation effect could cause the large ∼500K difference in temperature that is only seen in the central zone. Furthermore, no unusual LASER function can be detected. Hence, it is likely that the data reduction program is the cause.

3.8.3.2 Analysis

The data analysis program goes through an iteration process (section 3.5.4) to find a fit with the smallest χ^2 residual, when the measurement is compared against a modelled Gaussian convolved with the LASER instrumentation function. Initial parameters are set for the four output parameters, which are the background continuum intensity, peak intensity, peak position (i.e. LOS wind) and the width of the profile (i.e. temperature). Each of the spectral profiles from the 25 zones is fitted individually, and the initial parameters are only fixed at pre-set values at the beginning of a 25-zone fit. The final iterated results from each zone are passed on as the initial parameters for the next zone, except for the intensity background value. The iteration step size is approximately 0.2 percent of values of the parameters and the parameters are adjusted simultaneously. The limit of the number of iteration set to a high number at 1000 (section 3.5.4).

The zenith zone is coincidentally the first zone in a data set. It is the only profile in a 25-zone group that has pre-set initial parameters. The iteration process is printed out with the intermediate values of each of the parameters, and are shown in figure 3.15.

It can be seen that even though the first initial parameters are set to reasonable values at intensity background (Bg) = 100, intensity (In) = 1500, wind (Un) = -5 and temperature (Tn) = 750, results from the first iteration give a negative intensity background, and the fitted temperatures are subsequently falsely lowered to ∼200K for the zenith zone. The final fitted parameters at the end of the iteration are then passed on as initial parameters for the curve fitting in zone two, apart from the background value. The temperature results then return to normal after a few iteration loops for zone 2.

The IDL algorithm for the analysis performs a linear fit on the intensity background for every single iteration. Also if unreasonable results are found from the iteration i.e. the conditions that the background intensity is less than zero, and temperature less than 0.01K

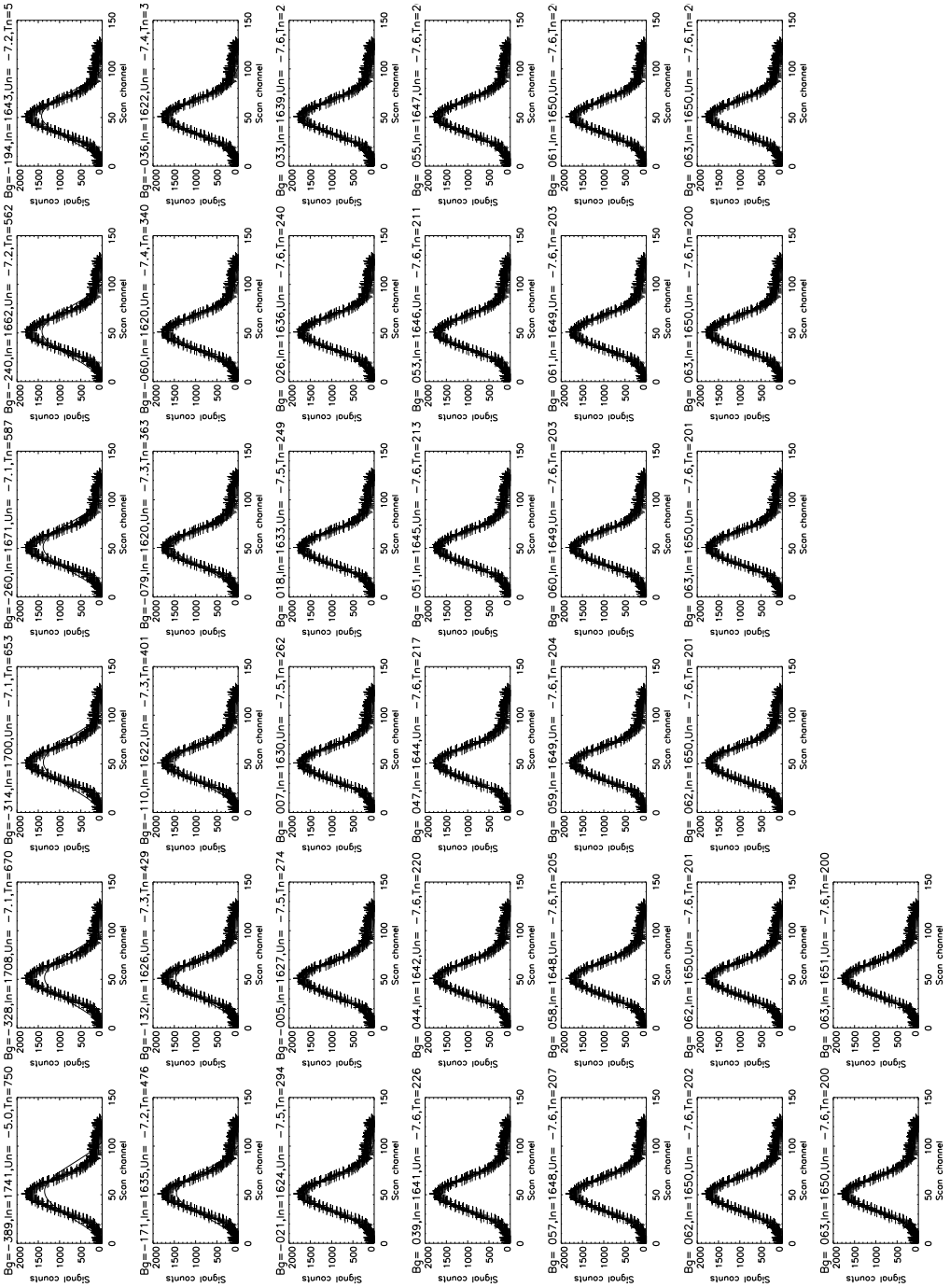


Figure 3.15: This figure shows the individual iteration profiles through the fitting process. The line graph represent the fit and the points represent the measured spectra. Each plot is labelled with the fitted background intensity, the peak intensity, the LOS wind and temperature at a point of the iteration process. These are all results from the zenith SCANDI zone, and the iterated results start from the top row, increases across the row, and then to the next row etc.

or greater than 3000K are met; the algorithm invokes a 'warp' by shifting the wind and temperature results artificially to avoid a bad local minimum. The negative background intensity from the linear fit for the zenith zone satisfies the 'warp' condition, and hence causes the temperature to be shifted to an unreasonable ∼200K.

3.8.3.3 Solution

Overall, since only the first fitted zone is affected, this suggests that the algorithms are very sensitive to the initial parameters. The simplest solution is to repeat the first zenith zone fitting once, discarding the first set of fitted results with the incorrect temperature, and pass on the final parameters of the first fit as initial parameters in the repeat of the first zenith zone fitting. This solves the central zone problem.

For the current algorithm, there is not yet any catch in the code to separate a bad minimum from a good one. However, it is possible to include a simple threshold limit into the code to avoid a bad minimum. At the same time, the method described in this thesis, where the first temperature derived by the analysis program is re-analysed to give a sensible set of initial parameters, is also able to help to avoid the 'bad minimum' issue.

3.8.4 Temperature ring offset and etalon parallelism

The other possible instrumentation effect is the temperature offset between different SCANDI rings, as demonstrated by figure 3.16.

3.8.4.1 Diagnostic

For the SKY data, there is a small approximate 20K average temperature difference between the different SCANDI rings, for the period December 2007 to October 2008. Please note the average temperature does not vary linearly from the zenith zone to the outer ring. At the beginning of season 2008/2009 for October 2008, the average temperature difference between the SCANDI rings increased significantly to approximately 500K.

For both seasons, the average temperatures vary as follow (in increasing temperature order): middle ring, inner ring, outer ring, and the highest temperature is found in the zenith zone.

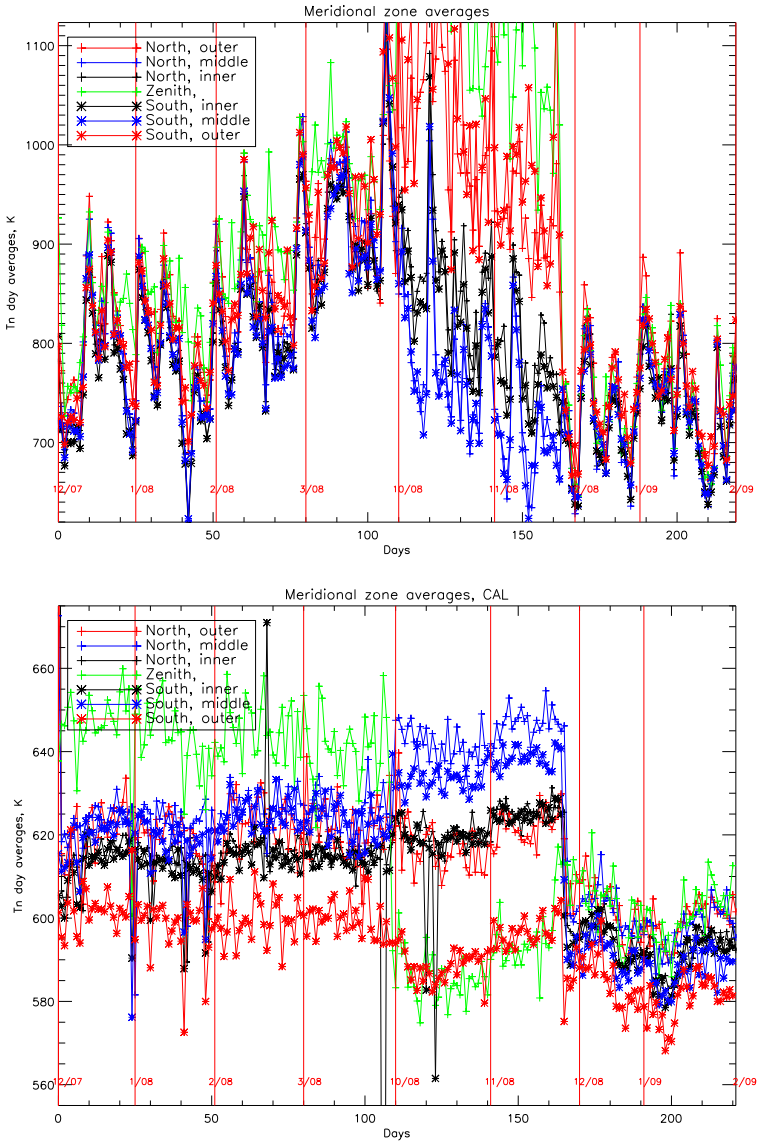


Figure 3.16: The two plots are similar, both showing daily average neutral temperature from December 2007 to February 2009. The difference is the top plot is actual SKY data, while the bottom plot is the calibration lamp CAL results. Only the meridional zone results are plotted. Graphs of the same colour are data from the same SCANDI ring. Red are data from the outer ring, blue from middle ring, black from inner ring, and green is the zenith zone. The vertical lines separate data from each month, and the month is written in red. Note that the SKY and CAL plots have different scales.

3.8.4.2 Analysis

It has been discovered by McWhirter, I. (private communication) that the temperature offset is caused by a slightly off-parallel etalon. In the phase mapping process (section 3.4), each pixel in the detector measures an interferogram. Each interferogram also peaks at different points of an etalon scan and mirror separation. The interferogram from each pixel is then 'phase-shifted' and summed together according to the zone it belongs to (section 3.4).

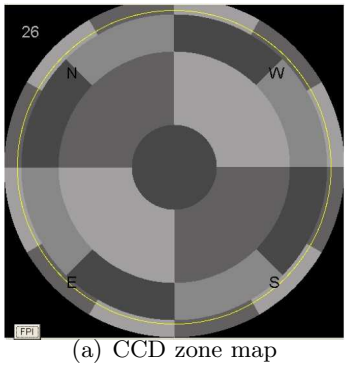
The main reason for the ring offset in temperature is that pixels that are associated with an interferogram which peaks at the 'more parallel' end of a scan contribute to a narrower average interferogram peak within a zone. In contrast, pixels from the 'less parallel' end of the scan contribute to a wider average interferogram within a zone.

If the zone division on the CCD (figure 3.17(a)) is compared against the phase map (figures 3.17(b) and 3.17(c)) where black indicates pixels that peak at the low end of the scan, and white at the high end of the scan, it can be seen that:

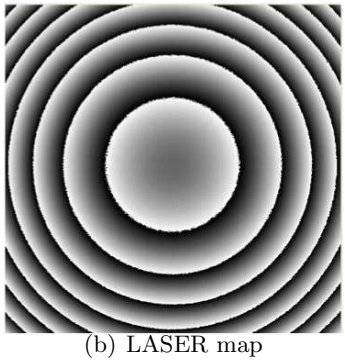
- The zenith zone has the majority of the pixels peaking in the same place, so it is heavily biased.
- The pixels in zones of the middle ring are distributed throughout roughly a single order, so it contains a good average.
- The pixels in the inner ring have about 1.5 orders and so have a bias to one end of the scan.
- The outer zone is roughly half an order, and is also biased to one end of the scan.

The different phase maps can be compared against each other. Contrasting the LASER phase map and the SKY phase map, their shadings are approximately inverse to each other i.e. the LASER phase map is approximately out of phase with the SKY phase map.

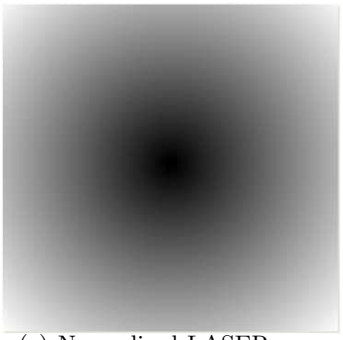
The FWHM of the spectral profiles (temperatures) are deduced from the deconvolution of the LASER function from the SKY function. The LASER instrument function for the zenith zone largely consists of signal recorded at the lower and more parallel end of the scan. In contrast, the SKY signal for the zenith largely consists of signal recorded at the higher (less parallel) end of the scan. This therefore leads to higher analysed SKY temperatures in the zenith zone, as shown in figure 3.16. The same argument can be applied to explain the other ring offsets in temperature.



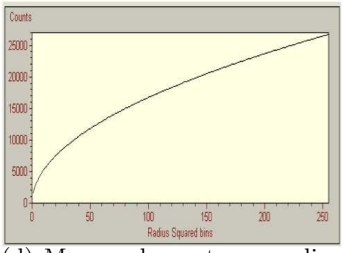
(a) CCD zone map



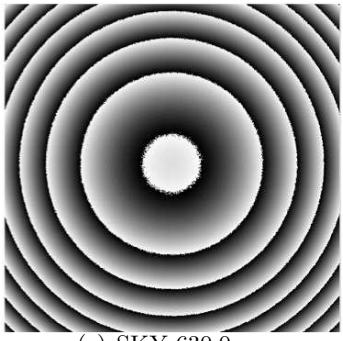
(b) LASER map



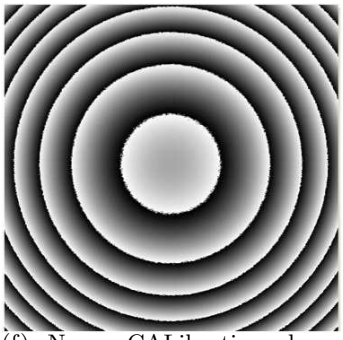
(c) Normalised LASER map



(d) Measured counts vs. radius square bin



(e) SKY 630.0nm



(f) Neon CALibration lamp 630.4nm

Figure 3.17: Courtesy McWhirter, I.. These are the figures showing the SCANDI zone division on the CCD that is appropriate for an equal area sky projection, a LASER phase map, a normalised LASER phase map, a r^2-bin reduced normalised LASER phase map plotted as counts vs. r^2 bin, and finally the phase maps for SKY and CAL.

As for the CAL results (figure 3.16), which has a much smaller ring offset than the SKY, the CAL is more or less in phase with the LASER phase map, hence the non-parallel etalon scan has a smaller effect.

3.8.4.3 Solution

The shift in etalon parallelism is caused by the analogue electronics used in the scanning, which can be prone to drift, even though it was designed carefully for maximum stability. Overall, the parallelism is maintained well on the whole, drifting remains in the differential gain between the three channels associated with the three legs of the etalon. SCANDI was fine-tuned and corrected by McWhirter, I. in 11/2008, and the ring offsets clearly disappeared from this date onwards, as shown in figure 3.16.

In figure 3.16, a jump in temperature is also seen in the CAL results. The jump is caused by the same on-site instrumentation maintenance, where the etalon scan was re-calibrated. As shown in this figure, the difference in temperature between the SCANDI zones was large between October 2008 and November 2008. The difference is then significantly reduced afterwards.

3.8.5 LOS wind

3.8.6 LOS wind relative spectral peak position and Doppler shift

The main calibration process for the LOS wind is to establish the correct baseline. The baseline is the instrumentation value that is added to the true wind data. In addition, SCANDI wind can be independently compared with the co-located Svalbard FPI, which has been producing wind data for 30 years at the time of this thesis.

3.8.6.1 Red and Blue shift directions for SCANDI and the FPI

For the single viewing FPI, the Doppler shift direction that is translated to the spectral r^2 space is simply determined by equation 3.13, where $\lambda \propto -r^2$ i.e. an increase in wavelength corresponds to a decrease in bin position in r^2 space. Hence red shift corresponds to a negative shift spectral peak position, and blue shift is represented by a positive shift.

However, for SCANDI, the process is slightly different because of the etalon stepping process. Consider a pixel at radius, r: when there is no red or blue shift, the peak response could be detected e.g. at etalon step 30. This is the zero velocity position. When there is a blue shift, the ring pattern expands similarly to the single view FPI. However, for

SCANDI, the peak response for the pixel is reached earlier in the scan e.g. step 28. The software then subtracts the zero velocity and gives, 28-30=-2. Finally, an empirical offset is added to the final value, so peaks do not centre around zero e.g. 60-2=58. This value can be artificially set to any value. For the current empirical offset value, blue shift is responded by a negative bin position shift, and red shift a positive shift. This is therefore opposite to the FPI.

3.8.6.2 FPI baseline

In terms of establishing the zero velocity baseline, the standard FPI technique assumes that the average zenith wind during a night of observation is zero. In addition, the calibration lamp data, which are taken with every complete directional cycle, are also used to account for any instrumentational drift.

For the FPI procedure, the average offset between the calibration lamp peak and the zenith peak is calculated for an individual night, then this offset is added to the 15 minutes interpolated calibration lamp peak data to produce a baseline.

3.8.6.3 Conde and Smith [1998] baseline

For Conde and Smith [1998], which describes an all-sky FPI similar to SCANDI, its baseline is established by first finding an average from the LOS wind data measured from each zone for each time stamp. Any outliers are excluded from the averaging. This produces a baseline data set consisting of the average wind data at each time stamp, which was measured at 15-minute time resolution at the time of the paper. It is then three-point smoothed to produce the final baseline.

The reasoning is that when considering a single all-sky measurement, it is possible to create a zero Doppler shift baseline by averaging the LOS winds from every sector within a single ring. This is because a ring-average of an uniform horizontal wind field should give a value of zero. This is illustrated simply by

$$U_{los}(i) = \underline{U} \cdot \underline{\widehat{r}}\,(i,j) \tag{3.33}$$

where U is the wind, U_{los} the LOS wind component seen by each zone and $\underline{\widehat{r}}\,(i,j)$ is the unit vector describing the look direction for the i-th zone in the j-th ring. The average around each ring hence give $\langle U_{los} \rangle = 0$.

This method assumes that any residual instrumentation effect is also evenly distributed around the zone map, and can be averaged out in a similar way. However, each zone might have a characteristic instrumentation contribution that cannot be accounted for by a single average over a zone map. So baselines for individual zones are investigated.

Also, the assumption that the zenith wind averaged to zero over each night might not be true, especially when the instrument is not covering a whole 24 hours of measurement as it is in mid-winter. So the effect of reduction in night time coverage on the baseline is also investigated.

3.8.6.4 Baseline test, daily average

The first test is to investigate the SKY and CAL daily-average data from each SCANDI zone, and results are shown in figures 3.18 and 3.19 . It can be seen that for all of winter 2007-2008, each zone has a significantly different baseline for the SKY data, while the CAL baselines are relatively similar apart from the baseline in zenith zone 1. Please note the vertical scale of the CAL plots are 100K, while the scale for the SKY data is 600K. Therefore, the CAL spikes are relatively small compared to the SKY data and do not show up in the final data.

The daily average results are expected to average to the baseline value. In addition to the reasoning provided by equation 3.33, the wind is expected to follow the anti-sunward direction, dominated by the solar-pressure force. Hence, the horizontal wind direction should cover a full 360^o rotation over a 24-hour period. This also explains the drift in wind value towards the end of the winter due to the reduction in observing hours.

Figure 3.20 summarises the number of data points collected at each hour and time zone. The difference in sample size for each hour is simply a reflection of the hours of darkness available for optical observation, apart from some very infrequent short periods of instrument maintenance.

The results here indicate that if daily average LOS wind data are involved in the baseline establishment process, it is possible that the baseline is biased towards LOS winds measured at particular hours that are not as affected by daylight. It can also be seen that the CAL fluctuations are at least a fifth smaller than the daily-averaged wind results. This indicates that the instrument itself is relatively stable in terms of spectral drift. About 3-4 CAL measurements are taken every night.

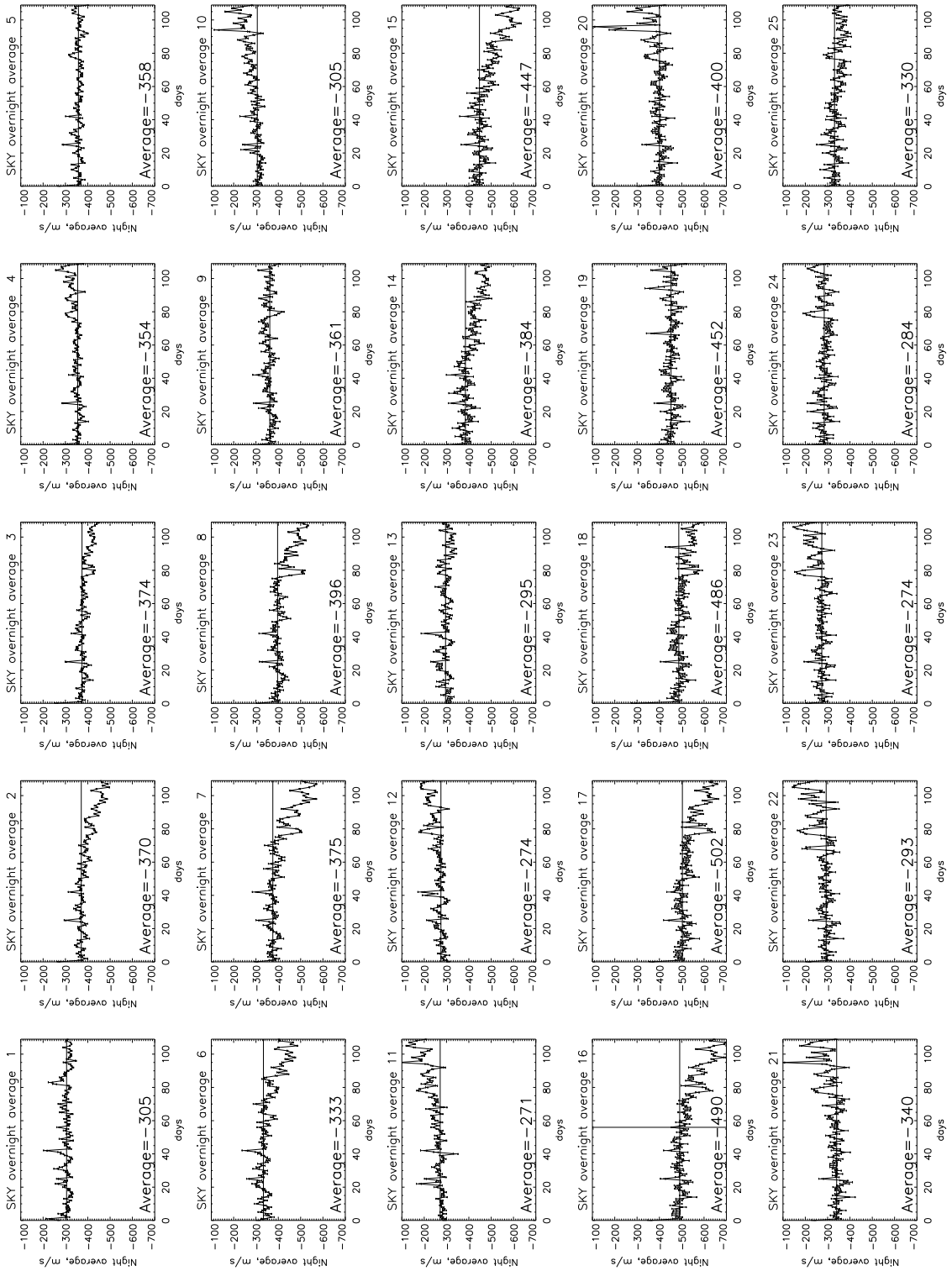


Figure 3.18: The daily average LOS wind throughout winter 2007-2008 are shown and each plot represents red-line data from a zone. These are measured SKY data. The average value is marked with a line with the value labelled. The wind values centred around about -300, which is the background value for data in the instrument before processing.

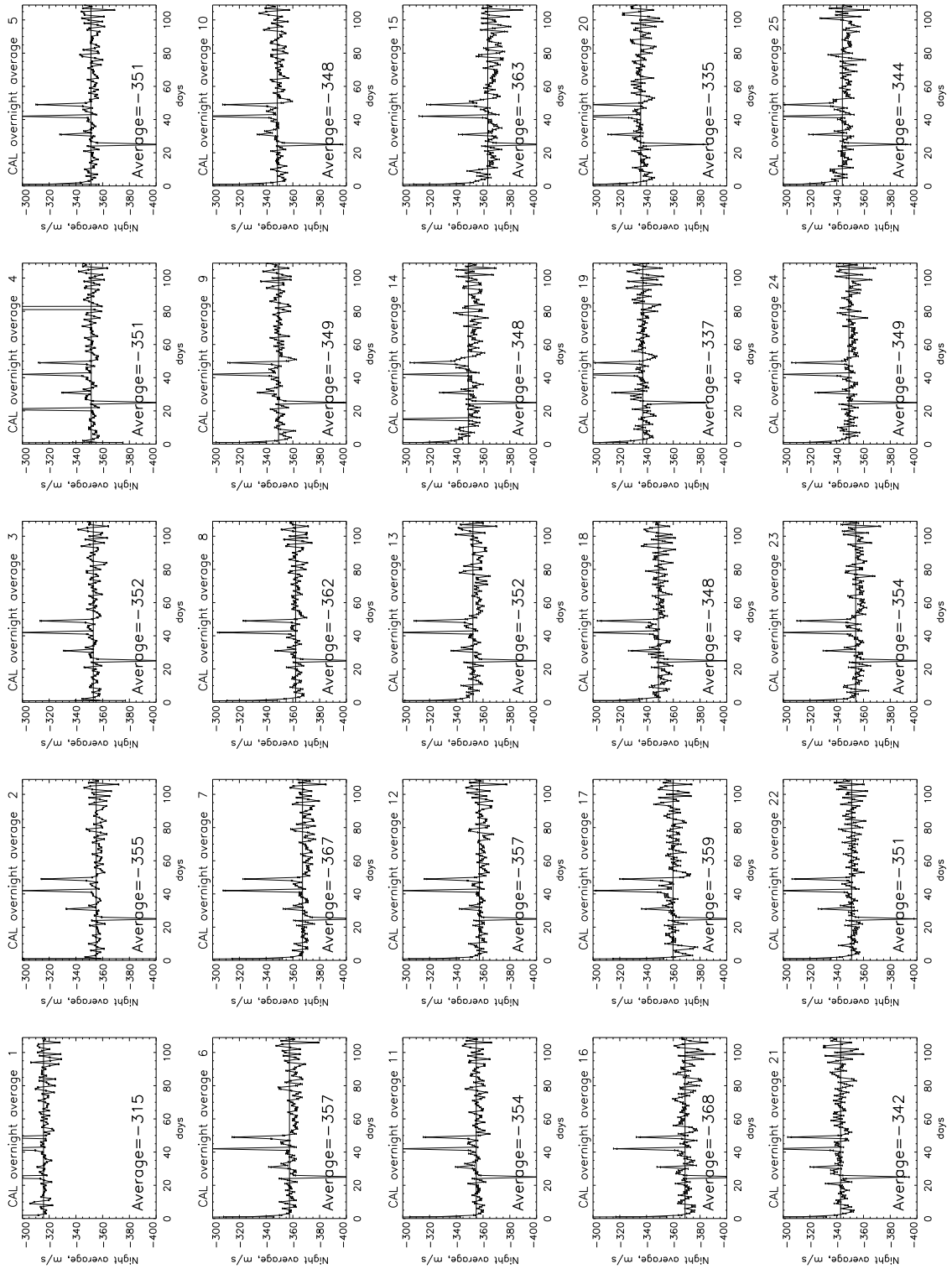


Figure 3.19: The daily average LOS wind throughout winter 2007-2008 are shown and each plot represents red-line data from a zone. These are CAL data. The average value is marked with a line with the value labelled. The wind values centred around about -300, which is the background value for data in the instrument before processing.

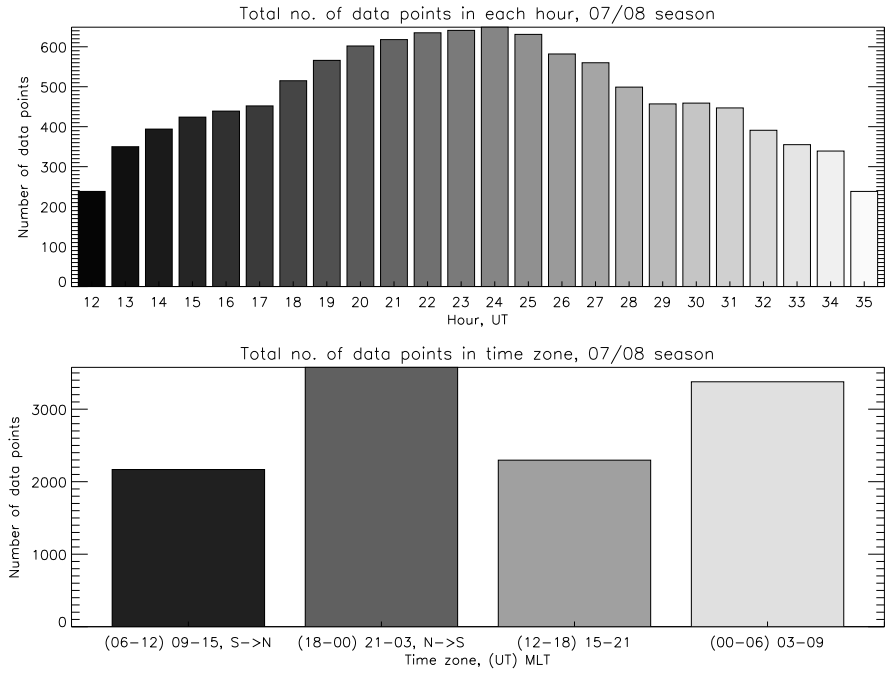


Figure 3.20: The top plot shows the number of data points for each measuring direction at different hour in UT, for winter 2007/08. The bottom plot shows the number of data points grouped into four 6-hour periods: 06-12, 19-00, 12-18 and 00-06 UT, or in MLT they are: 09-15, 21-03, 15-21 and 03-09.

3.8.6.5 Baseline test, season average

Another significant variation shown in figure 3.18 is the averaged value over the whole observing season. A mean is taken over the day-averages, excluding days that are significantly affected by daylight. Over the 25 zones, a minimum average of -502 m/s and a maximum of -271 m/s are detected. Again, the season average detected in the CAL is much smaller with a difference of $\sim$10m/s. It is expected that the season average should be a rough representation of the instrumentation baseline, and the results here seem to indicate that each zone unexpectedly has a different baseline. The season averages are plotted in figure 3.21.

Variations in the season average can be seen in both SKY and CAL results, and the SKY data clearly follow a sinusoidal change in the middle and outer rings. There are several possible sources that can cause the sine variations:

- It could be a reflection on a general LOS wind trend over the observing season. In this case, it would be a general south-east background flow according to the phase of the sine changes. However, this is unlikely because the wind pattern should be

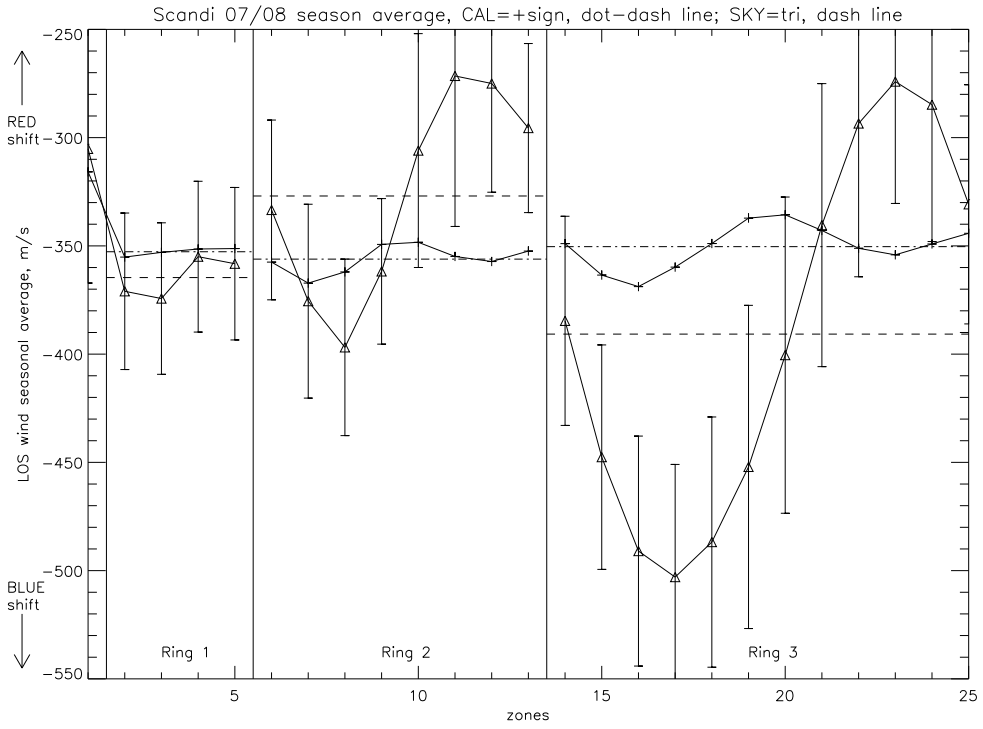


Figure 3.21: These are the red-line LOS wind season average for the 25 SCANDI zones. The triangle points are the SKY data, and the plus signs are CAL results. The error bars are the standard deviation of the data points (excluding outliers). The vertical lines separate the three SCANDI rings. The dot-dash horizontal lines are the season-ring-CAL average, and the dash lines are the season-ring-SKY average.

dominated by the solar-pressure force and follow a general anti-sunward direction, and hence averaged to ∼zero baseline level over a 24-hour period. Furthermore, a general anti-sunward flow is clearly seen in the horizontal wind field data (section 3.9) . Hence the sine variation is unlikely to be a measurement of physical effects.

- Another cause is the sampling. The previous point is only true if the data set covering the full 24-hours are used to determine the average. As can be seen from figure 3.20, the number of data points are not evenly distributed over all UTs, and there could be more hours in the sample which corresponds to the south-east UT.

- The data quality (cloudy/clear sky conditions) could also be the cause. However, it will be shown that excluding cloudy data does not cause significant change to the sine pattern.

- It could be a pure instrumentation effect e.g. lack of parallelism in the etalon

Finally, the green-line results from December 2008 are shown in figure 3.22, which have a similar sine pattern, but with a different phase. Hence, if the sine variation is an instrumentation effect, a different set of baselines need to be determined for the green data.

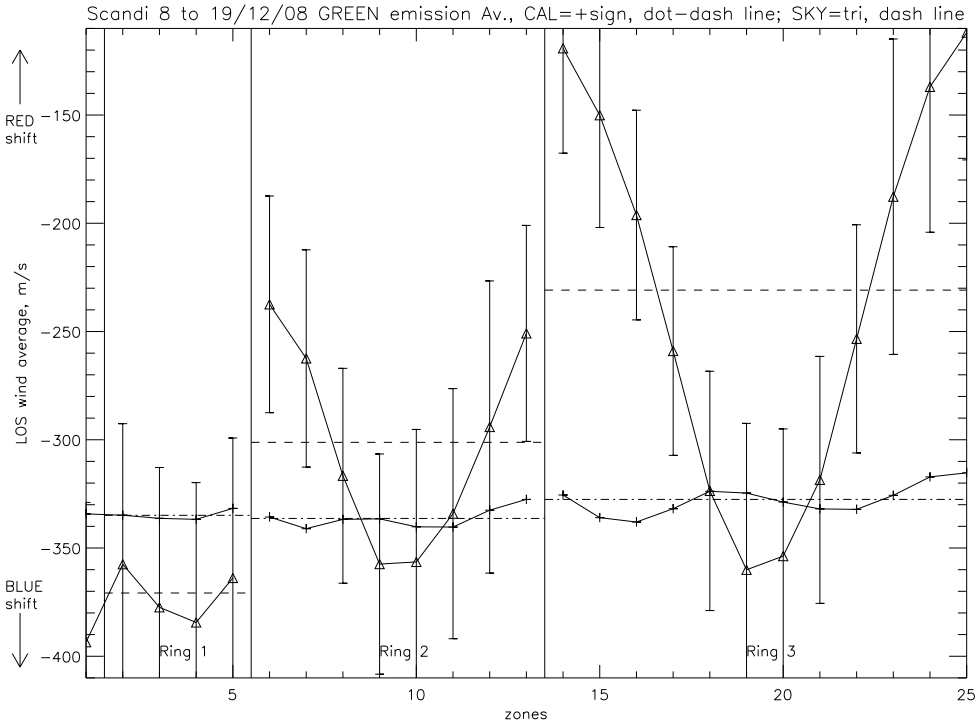


Figure 3.22: These are the LOS wind season average for the 25 SCANDI zones. This is similar to figure 3.21, but green-line data, instead of red-line data are shown. The data are averages from 8 to 19/12/2008. The triangle points are the SKY data, and the plus signs are CAL results. The error bars are the standard deviation of the data points (excluding outliers). The vertical lines separate the three SCANDI rings. The dot-dash horizontal lines are the season-ring-CAL average, and the dash lines are the season-ring-SKY average.

3.8.6.6 Baseline test, sampling

In order to discover the cause of the sine variation listed in section 3.8.6.5, different methods of sampling are tested. Figure 3.23 shows the season-averages of LOS wind measured from different zones.

The top and bottom plots in figure 3.23, i.e. the all-data sample and the clear-sky data only sample respectively, are compared against each other. The main difference is the clear-sky data only result gives a larger amplitude in the sine variation. This is expected because of cloud scatter in the LOS wind, which averages out any spatial information in

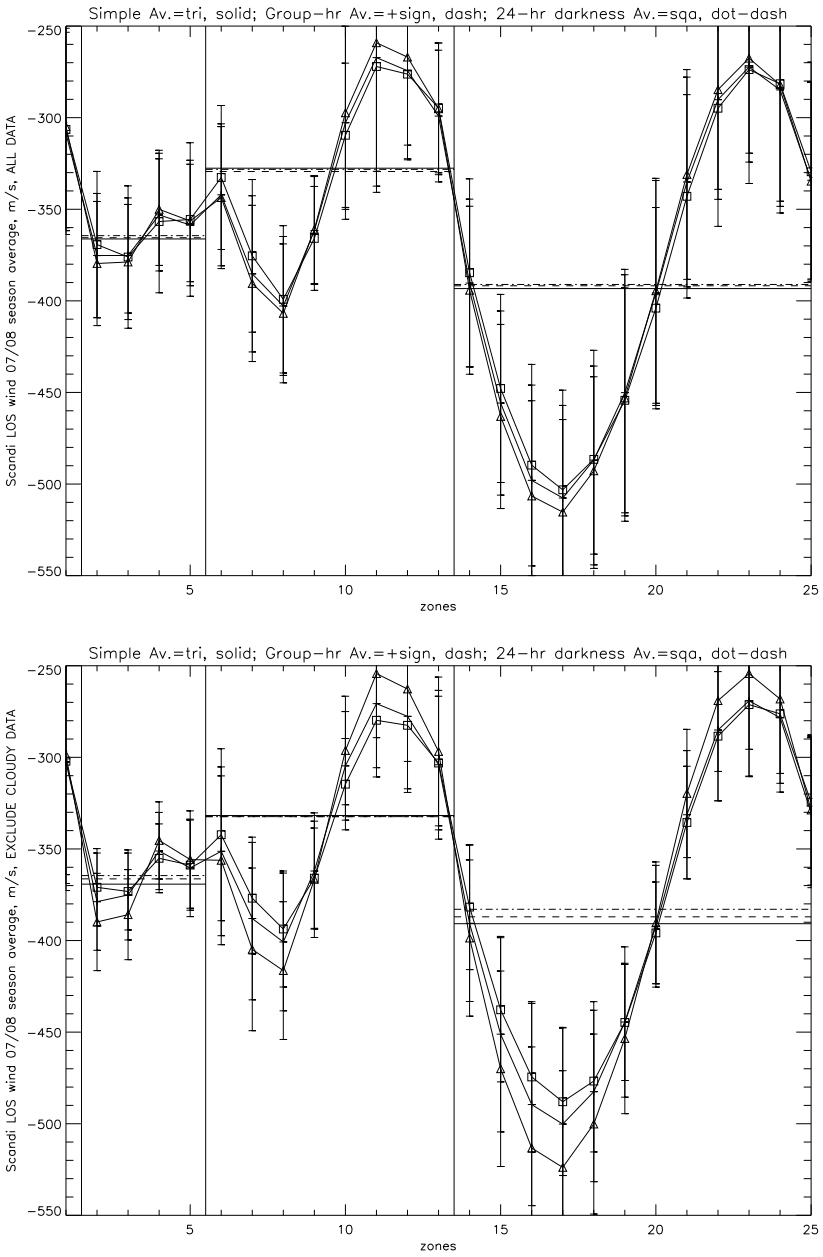


Figure 3.23: Each plot shows the red-line LOS wind season-average (y-axis) vs. 25 zones (x-axis). The vertical lines separate the 3 SCANDI rings, the horizontal lines are the ring averages, which are the mean of the zone averages in each ring. The error bars are the standard deviations. The only difference between the top and bottom plots is that the top plot includes all data, and the bottom includes clear-sky data only. The details of what the symbols and different plot lines represent can be found in table 3.5.

Table 3.5: A table describing the different graphs in figure 3.23. The information here is applicable to both plots in the figure, where the only difference between the plot is the top plot includes all data, and the bottom plot has clear-sky data only.

	Averaging Methods	Symbols in the zone Av.	Symbols in the ring Av.
1	Simple los wind average over the season	△	Solid line
2	Season average determined from first averaging over UT	+	Dash line
3	Averages from days of complete darkness only	*square*	Dot-dash line

the wind.

In terms of the three different methods listed in table 3.5 that are applied to the plots in figure 3.23, all methods follow a similar sine pattern, for both the all-data and clear-sky data only sample:

- Method 1 takes a simple LOS average over the season, and it gives the highest sine amplitude for both the all-data and clear-sky only sample. Its sine amplitude is higher in the clear-sky only case.
- Method 2 first averages the whole data set according to UT, before the season average is determined. This helps to eliminate some of the bias towards days which have less than 24 hours of data (figure 3.20). The sine amplitude is reduced by 50K comparing to method 1, indicating that non-24-hours sampling contribute partly to the sine variation.
- Method 3 includes only data when there are 24 hours of darkness. The results are similar to method 2.

To summarise, the test using the three methods above cautions against the use of non-24 hour samples to determine a baseline. Furthermore, cloudy data cause averaging in the LOS wind measurements because of cloud-scatter, and do not have any significant effect on the final baseline, and hence are not a source of the problem. None of the factors considered above reduces the sine pattern significantly. Hence the sine pattern is likely to be an instrumentation effect. A final test is to compare the different baseline methods with the FPI LOS data (section 3.8.6.8) .

3.8.6.7 Baseline test, etalon parallelism

A possible instrumentation effect on the LOS wind is the etalon parallelism (section 3.8.4). The fine-tuning of etalon parallelism has solved the problem of temperature offsets between different SCANDI rings, and it is possible that the sine variation is also an effect of the change in parallelism at the start and end of an etalon scan.

Figure 3.24 shows the season average results before and after the etalon tuning. It can be seen that although the sine amplitude is reduced, especially for the middle ring (ring 2), the sine pattern is still present in the average LOS wind results. This suggests that etalon parallelism is only partly responsible for the sine pattern. This is expected because the parallelism mostly affects the FWHMs of the LASER and SKY spectral profiles, and the effect on the peak position of the profiles can therefore remain small.

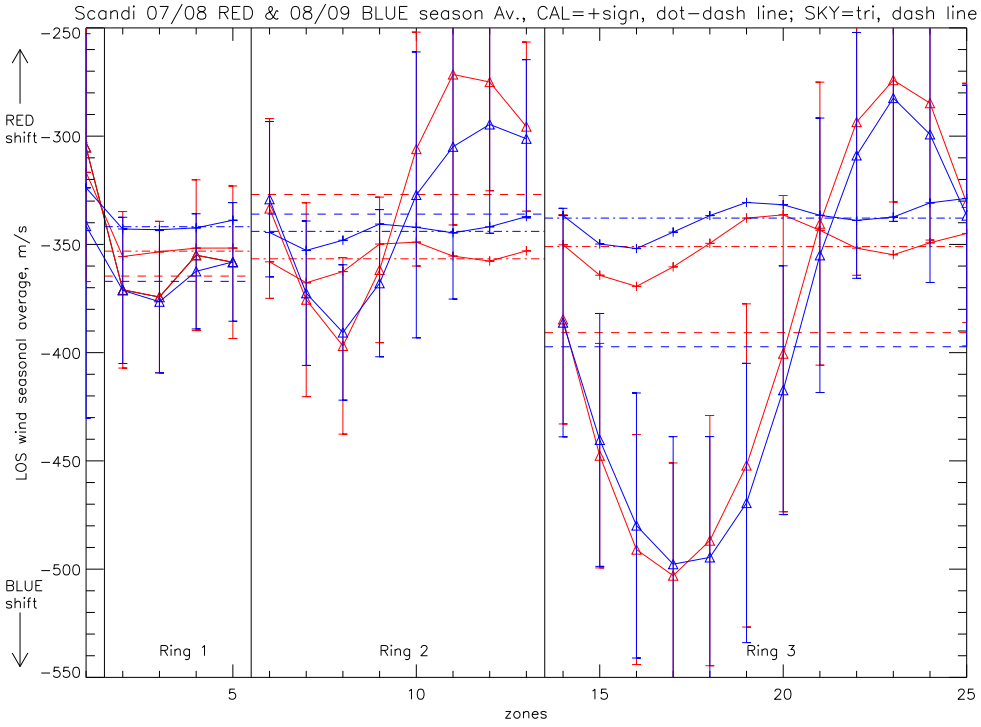


Figure 3.24: These are the LOS wind season average for the 25 SCANDI zones. This is similar to figure 3.21, but compares the results before (red) and after (blue) the etalon parallelism fine-tuning. The red data are from season 2007/08 and the blue data are from season 2008/09. The triangle points are the SKY data, and the plus signs are CAL results. The error bars are the standard deviation of the data points (excluding outliers). The vertical lines separate the three SCANDI rings. The dot-dash horizontal lines are the season-ring-CAL average, and the dash lines are the season-ring-SKY average.

3.8.6.8 Baseline test, SCANDI-FPI comparison, method 1

So far the sine variation still cannot be explained fully, by both the data sampling methods, as well as instrumentation effects. The FPI data are compared against the SCANDI data as quality check.

SCANDI data from the North, South, East and West zones in the outer ring are compared with the co-located FPI wind results. SCANDI outer ring has the largest amplitude in the sine variation, and its comparison with the FPI results should give a good indication of whether the sine pattern is due to an instrumentation effect. The SCANDI and FPI data are measured independently. The FPI baselines are determined from the overall effect of the average zenith wind over a night and the calibration peak fluctuations (section 3.8.6.2). To avoid any bias in the baselines from non-24 hour measurements in the zenith wind contributions, December data are chosen.

First, the data set from the two instruments are directly compared against each other, and are shown in figure 3.25. For this case, the SCANDI baseline is simply the outer ring average. The SCANDI data plotted here are taken from the north, east, south, west zones within the outer ring. The same outer ring average is used for the four zones. This gives an indication of the effect of applying a single baseline for the SCANDI data.

The results show a distinctive systematic offset between the FPI and SCANDI results. The closest match between the two instruments is the north direction, and the FPI LOS wind is lower than SCANDI by ∼100 m/s in the east direction, ∼80 m/s in the south and west. This indicates that a single baseline for all SCANDI zones is not the appropriate method, assuming the FPI data are accurate. When cloudy and noisy data are taken out of both the FPI and SCANDI data, the results remain similar and the data quality is not a source of the difference.

3.8.6.9 Baseline test, SCANDI-FPI comparison, method 2

A different SCANDI baseline is tested next: the baseline is determined individually for each zone, from the zone averages over a 24-hour night. The results are shown in figure 3.26. It can be seen that the offsets between the FPI-SCANDI data are significantly reduced, though a smaller difference of ∼40m/s can still be seen in some nights/directions. This seems to indicate that the zone-average baseline is more appropriate for the 2007/08 data. The data set is investigated further with correlation tests between the FPI and SCANDI samples.

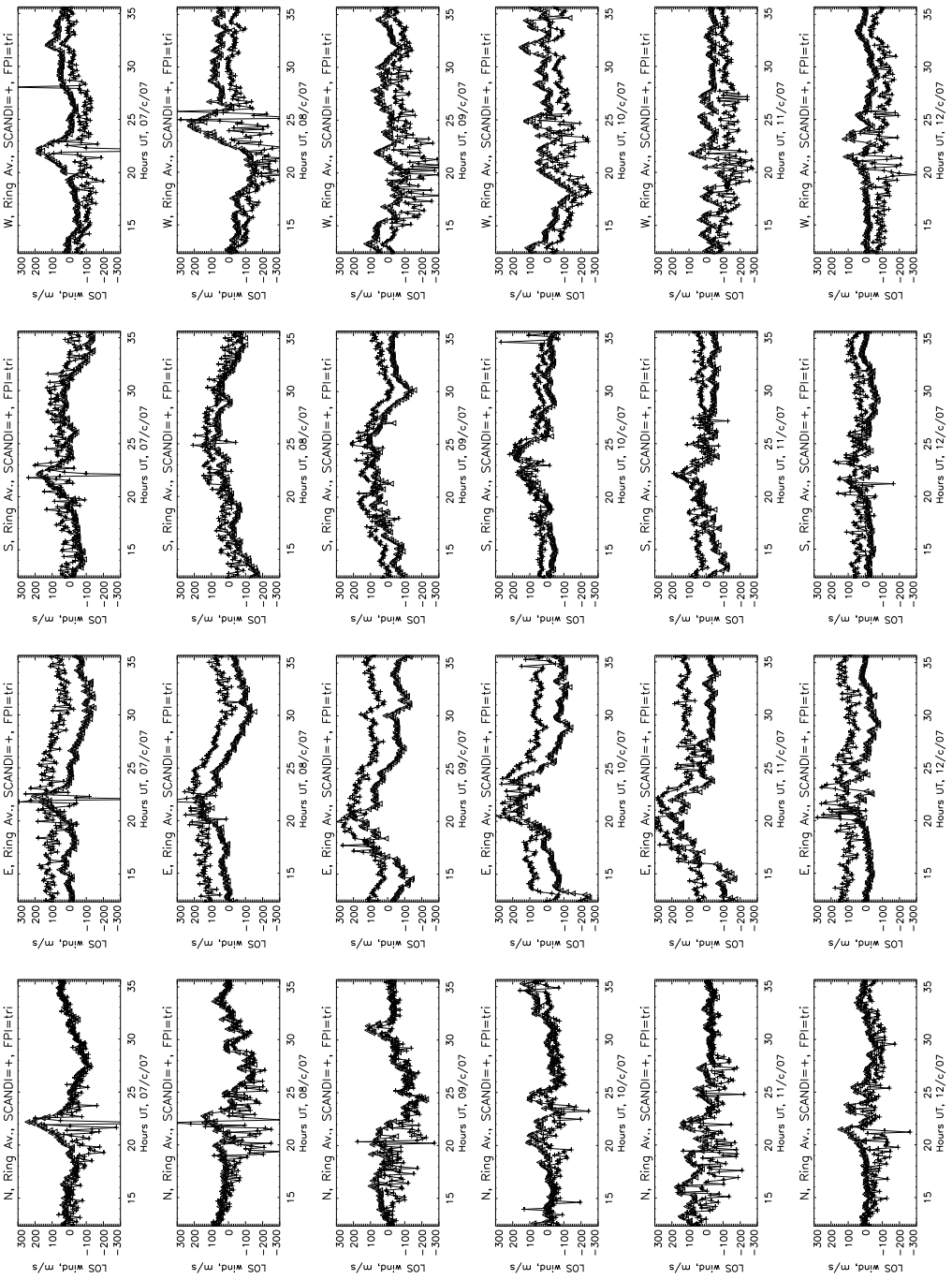


Figure 3.25: Each row in this figure shows data from one night of FPI and SCANDI data, and the columns are data from the north, east, south and west directions for FPI 1 degree FOV LOS wind (triangles), and the associated SCANDI zones in the outer ring (plus signs). The baselines for the FPI data are determined using a method which considers the zenith wind average and CAL variations. The SCANDI baselines in this display are determined from the average wind values from the SCANDI zones in the outer ring .

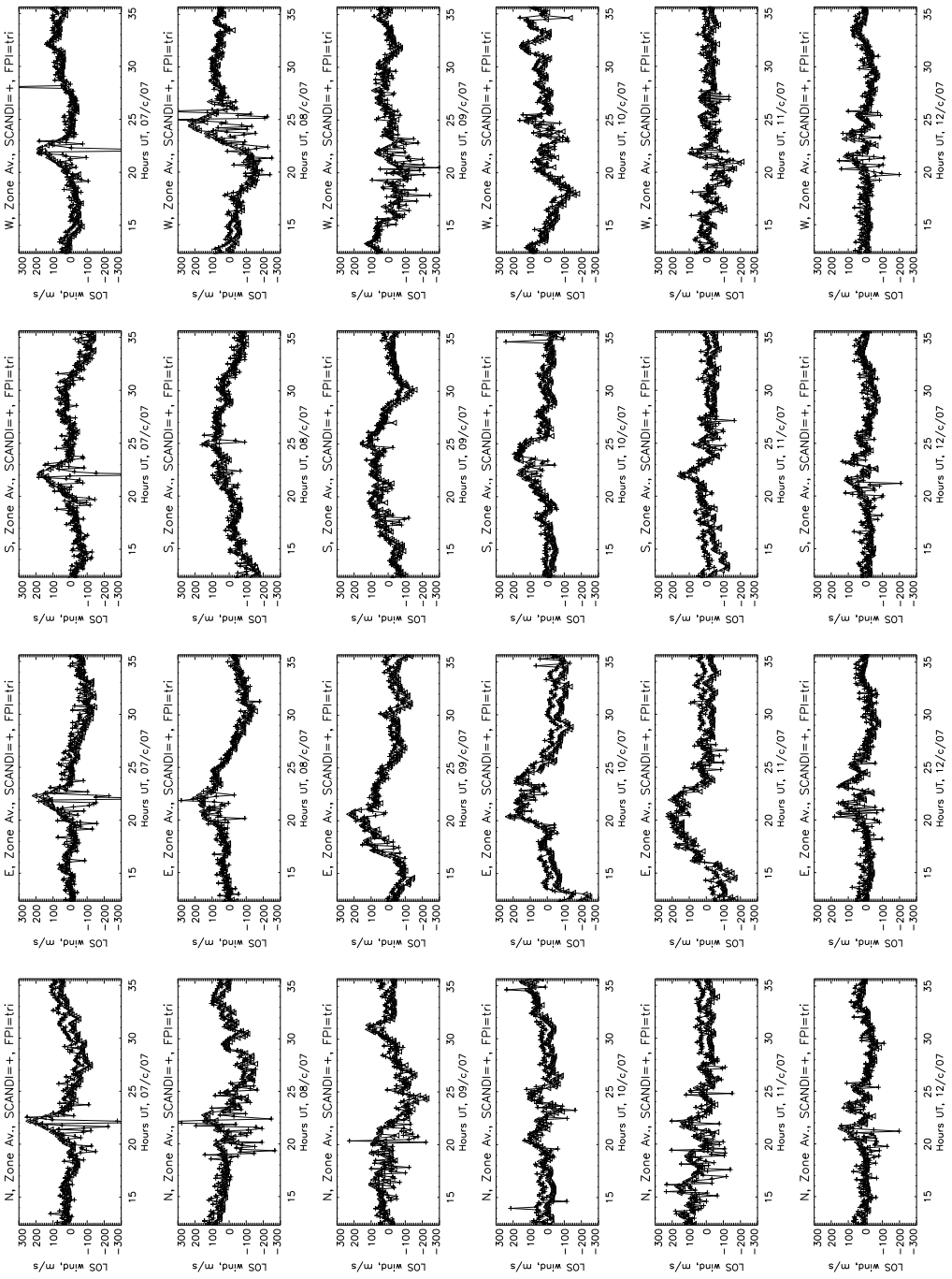


Figure 3.26: This figure is similar to figure 3.25, and the only difference is that a different baseline is used for each zone, calculated from the zone's night average. Each row in this figure shows data from one night of FPI and SCANDI data, and the columns are data from the north, east, south and west directions for FPI 1 degree FOV LOS wind (triangles), and the equivalent SCANDI zones in the outer ring (plus signs). The FPI data have the usual baseline determined from the zenith wind average and CAL variations.

3.8.6.10 Baseline test, SCANDI-FPI comparison, correlation over one night period

The exact cause of the systematic difference has not yet been found at the time of this thesis. It could be an artifact from e.g. the etalon scan and lack of parallelism. However, the SCANDI temperatures after the baseline correction match with the independently measured FPI data, which give confidence that the SCANDI data are correctly analysed.

Following the conclusion that the individual zone average gives the best match baseline between SCANDI and the FPI, the next step is to decide if the single night average is a suitable baseline, or if a time series of average data as it is used in Conde and Smith [1998] is more suitable (section 3.8.6.3). Therefore the correlation of the time series of LOS wind data is computed.

It should be noted that the time series average method cannot be applied with method 2 (section 3.8.6.9), which requires a different baseline for each SCANDI zone. For a time series baseline to be valid, the value of the baseline at each time stamp needs to be independent from the natural atmospheric variations of the neutral wind. Hence, the time series average method requires equation 3.33 to be satisfied. In this case, the data is averaged around the 360 degrees azimuth, creating a 'ring average'. Please note the 'ring average' is a tool for creating a time series in the correlation test and is not expected to replace the preferred baseline method 2. The correlation results are shown in figure 3.27.

The results show that the time-independent baseline determined from a single-night gives the highest number of nights with best correlation between the FPI and SCANDI data. This is different from the three-point smooth method (on 15 minutes resolution data) chosen by Conde and Smith [1998].

Nevertheless, the time-independent baseline determined from a single overnight average provides the best correlation between SCANDI and FPI data. This also reassures the validity of the preferred baseline method 2, which needs a different baseline for each SCANDI zone and is only valid for the time-independent baseline method. It also indicates the instrument is very stable and there is negligible spectral drift over each night. Hence a time-independent baseline is all that is needed.

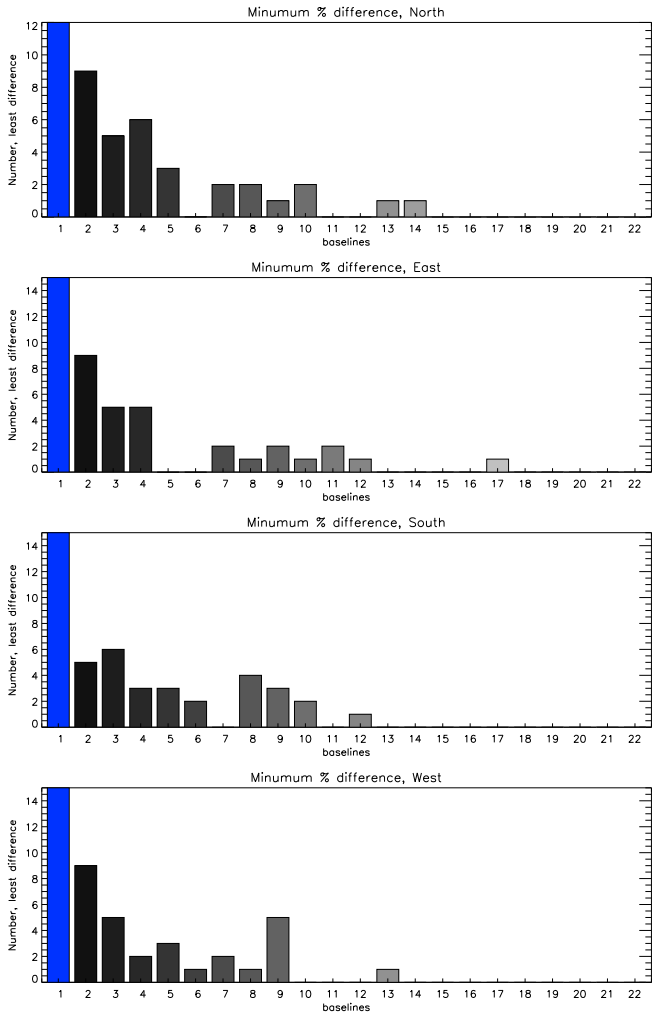


Figure 3.27: These are the correlation results between SCANDI and FPI LOS wind data from season 2007/08. The SCANDI baselines are determined from the average wind value from zones within a SCANDI ring. Each plot represents a correlation test at different directions from North, East, South and West from top to bottom. The numbers from 1 to 22 on the x-axis represent different SCANDI 'ring average' time series baseline. 1=baseline with a single value calculated from the overnight average. 2- 22=a time varying baseline with different degree of smoothing. 2=single point/no smoothing, 3=three-points smoothed,..., 22=forty-one-points smoothed. The time resolution of SCANDI data is ∼7 minutes. The y-axis indicates the number of nights that give the best SCANDI-FPI correlation results for each baseline method.

3.8.6.11 Baseline test, SCANDI-FPI comparison, correlation over longer period

The next step is to test the baseline average for even longer periods of time, determined from week-average, month-average and season-average. The results are shown in figure 3.28.

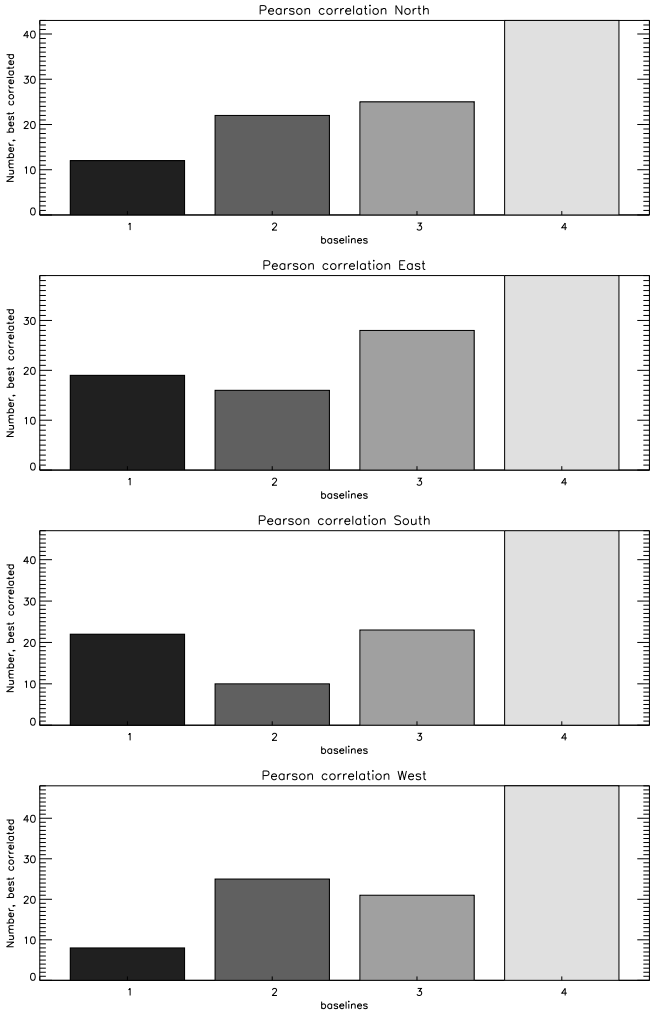


Figure 3.28: This is similar to figure 3.27, but the baseline method is different. These are the correlation results between SCANDI and FPI LOS wind data from season 2007/08 with different SCANDI baselines. Each plot represents correlation test at different directions at North, East, South and West from top to bottom. The numbers from 1 to 4 on the x-axis represent different SCANDI baselines. 1=single outer ring overnight average baseline, 2=single outer ring baseline average over a week of data, 3=average over a month, 4=average over a season. The y-axis is the number of nights that gives the best SCANDI-FPI correlation results from each of the baseline method.

They show that the best correlated method comes from a single baseline calculated from the whole season of data. This is again an indication that there is very little instrumentation drift in the SCANDI data, and the time-series method may introduce unnecessary variations in the baseline. Method 2, the individual zone average baseline, can therefore be applied.

3.8.6.12 Baseline test, conclusion

- A definite sine variation vs. SCANDI zones in a ring can be seen in the season-average data, for both red and green-line emissions, with a slightly different phase.
- Non-24-hours sampling contribute partly to the sine pattern, because only part of the detected 360^o rotation of the general anti-sunward flow would be included in the sample.
- Quality of data set, i.e. cloud and noise, is not a source.
- Fine-tuning of etalon parallelism has partly reduced the sine pattern amplitude, but does not fully explain it.
- There is no obvious physical effect that can cause a general north-south background flow in the red-line data over a whole observing season, with a near perfect sine variation vs. zones. The SCANDI data are a good match to FPI data once the baseline variations are accounted for with the correct method.
- Correlation tests between FPI and SCANDI data show that a smoothed time series from the average LOS wind in the observing zones [Conde and Smith, 1998] may introduce unnecessary variation in the baseline for SCANDI.
- For the 2007/08 data set, the best match, in terms of both correlation and absolute values, is found when each SCANDI zone has a different baseline that is determined from the zone-dependent season-average LOS wind value.
- In practice, daily zone-averages are calculated for each zone. They are then used as the zone-dependent baselines for immediate initial analysis of each night's data during the observing season. The data are then carefully monitored to avoid incorrect baselines.

3.9 Horizontal wind field

3.9.1 Introduction and background

After the correct baselines for the LOS wind are established, the next step is to determine the horizontal wind vector field. Both the IDL algorithms and theories involved are based on Conde and Smith [1998]

The horizontal vector fitting method originates from analysis devised for tropospheric Doppler radar data by Browning and Wexler [1968], which was then adapted by Burnside et al. [1981] for measurements in thermospheric airglow, and extended to allow estimation of total vorticity. The analysis of interest here is the deduction of horizontal wind component normal to our instrument's LOS.

It is not possible to deduce all three components of a measured wind field unambiguously from a single station line-of-sight data only. Tri-static multi-site experiments are required for a full solution. However, the following two assumptions related to the vertical wind can be applied in order to determine the vector field that could produce the LOS wind observed: the vertical wind is constant across the entire region of the instrument's FOV, and that the vertical winds are normally small compared to the typical horizontal wind of hundreds of m/s. Even though non-uniform vertical wind field can be present, it should remain relatively small compared to the horizontal wind magnitude. Furthermore, the vertical component in the measured LOS wind also decreases with increasing zenith angle. Therefore, distortion due to the vertical wind is not expected to be major. Its effect is investigated for calibration purposes.

3.9.2 Theory and method

The total wind field is resolved into two contributions, which are the scalar vertical wind and a two-component vector horizontal wind. Consider SCANDI, which samples spectra from sets of annular regions, i.e. rings, in the sky, where each ring is centred around the zenith, with a radius specified by zenith angle, ϕ. Each ring is then divided in azimuth into n equal-sized sectors i.e. zones, with the k^{th} zone centred on azimuth angle, θ_k.

For a given ring at zenith angle, ϕ, the set of measured LOS wind from each zone in the ring are all used to derive the estimation of wind vectors. The set of LOS wind in a ring is denoted as, $\left\{V^{\parallel}\left(\theta_k, \phi\right)\right\}$. $\parallel$ denotes that the velocity component is parallel to the LOS of the instrument observing zone.

One of the vertical wind assumptions is applied, where an uniform vertical wind, V_z, can be related to the LOS wind as

$$V^{\parallel}\left(\theta_k, \phi\right) = V_z cos\phi. \tag{3.34}$$

However, results from more than one ring are need. The reason is to distinguish a uniform vertical wind field, from a uniform divergent horizontal wind, ε, receding at the

same speed in all zones in the ring, represented by $V^{\parallel}(\theta_k, \phi) = \varepsilon cos\phi$. This is done by exploiting different $V^{\parallel}(\theta_k, \phi)$ dependencies on ϕ for a uniform vertical (V_z) and a uniform diverging wind field (ε).

V_z is directly measured by the zenith central zone, with ϕ=0 and n=1:

$$V_z = V^{\parallel}(\theta_1, 0)\,. \tag{3.35}$$

The uniform vertical wind field assumption allows the vertical wind contribution to be removed from the measured LOS wind from each zone. The convention here is positive away, or upward for V_z. The 'corrected' LOS wind that contains only the horizontal component, $H^{\parallel}(\theta_k, \phi)$, for the k^{th} zone is therefore described by

$$H^{\parallel}(\theta_k, \phi) = V^{\parallel}(\theta_k, \phi) - V_z cos\phi. \tag{3.36}$$

$H^{\parallel}(\theta_k, \phi)$, with the convention of receding wind being positive, can also be described in terms of the spatially varying zonal, H_x, and meridional, H_y, wind fields:

$$H^{\parallel}(\theta_k, \phi) = sin\phi\,(H_x sin\theta_k + H_y cos\theta_k)\,. \tag{3.37}$$

H_x and H_y can be expanded with the first-order Taylor series i.e. the zonal and meridional components about the zenith, can be written as:

$$H_x = u_0 + \frac{\partial u}{\partial x}x + \frac{\partial u}{\partial y}y \tag{3.38}$$

$$H_y = v_0 + \frac{\partial v}{\partial x}x + \frac{\partial v}{\partial y}y \tag{3.39}$$

where x and y are the zonal and meridional distances from the zenith to a zone, at zenith angle, ϕ, and azimuth angle, θ. Note that x and y can be written as $x = R\ sin\theta$ and $y = R\ cos\theta$ where $R = h\ tan\phi$ and h is the height of the emission layer.

The horizontal component of the LOS wind (equation 3.37) can also be represented as a Fourier series by

$$H^{\parallel}(\theta_k, \phi) = \left\{ a_0 + \sum_{m=1}^{\frac{n}{2}-1} (a_m sin\ m\theta_k + b_m cos\ m\theta_k) \right\} sin\ \phi. \tag{3.40}$$

It can be inverted to give the Fourier coefficients:

$$a_0(\phi) = \frac{1}{2n\ sin\phi} \sum_{k=0}^{n-1} H^{\|}(\theta_k, \phi) \tag{3.41}$$

$$a_m(\phi) = \frac{1}{n\ sin\phi} \sum_{k=0}^{n-1} H^{\|}(\theta_k, \phi)\ sin\ m\theta_k \tag{3.42}$$

$$b_m(\phi) = \frac{1}{n\ sin\phi} \sum_{k=0}^{n-1} H^{\|}(\theta_k, \phi)\ cos\ m\theta_k. \tag{3.43}$$

with the coefficients 3.42 and 3.43 evaluated for m=1 to m=n/2-1.

Conde and Smith [1998], Browning and Wexler [1968] and Burnside et al. [1981] applied the inter-relationship between equations 3.37 to 3.43 to determine the unknown terms, including the partial derivatives, in equations 3.38 and 3.39.

The Fourier coefficients for m=1 can be related to terms in equations 3.38 and 3.39, as (e.g. Browning and Wexler [1968]):

$$u_0 = a_1 \tag{3.44}$$

and

$$v_0 = b_1. \tag{3.45}$$

The remaining terms, i.e. four partial derivatives, are related to m=0 and m=2 coefficients, but only one nonzero m=0 coefficient is available. Hence, one of the partial derivatives, $\frac{\partial v}{\partial x}$, needs to be determined through other means for a unique solution.

The assumption is also based on methods by Browning and Wexler [1968] and Burnside et al. [1981], and it requires data from several measurements from consecutive time stamps. It is assumed that for short time intervals,

$$\frac{\partial v}{\partial x} \simeq \frac{1}{\epsilon} \frac{\partial v}{\partial t} \tag{3.46}$$

where ϵ is the tangential velocity of the observatory due to the rotation of the Earth. The assumption is that for sufficiently short time intervals, the observing station can be seen as moving though a meridional wind field that is stationary in local time, as the station moves with the rotation of the Earth. So the meridional wind, v, can be sampled at various locations along the zonal direction, x.

Conde and Smith [1998] cautioned that this assumption can be invalid when there is a sudden change in IMF B_y, since variations in IMF change the ionosphere, and hence the

thermospheric neutral circulation. The wind field seen in latitude/local time coordinates may not be stationary under these conditions.

The time resolution at the time of the Conde and Smith [1998] paper was 15 minutes, and resolution of SCANDI measurements is higher at 7 minutes. Therefore the assumption is more likely to be valid for SCANDI data because of the shorter time between exposures. The time constant for the neutral thermosphere to respond to a change in ion drag forcing is typically an hour or more, or at ∼50 minutes under high electron density regions like the cusp or evening auroral zone (Conde and Smith [1998]). This is clearly much longer than the 14 minutes span of two SCANDI measurements. $\frac{\partial v}{\partial x}$ is also small compared to the other derivatives.

In practice, $\frac{\partial v}{\partial t}$ from equation 3.46 is determined by the central difference approximation:

$$\frac{\partial v(t_j)}{\partial t} \simeq \frac{v_0(t_{j+1}) - v_0(t_{j-1})}{t_{j+1} - t_{j-1}} \tag{3.47}$$

where $\frac{\partial v(t_j)}{\partial t}$ is $\frac{\partial v}{\partial t}$ at the j^{th} time interval t_j, estimated from the uniform component of the meridional wind as a function of time, $v_0(t_j)$, which is three-point smoothed temporally before being evaluated by equation 3.47.

Finally, the following describes the relationship of the three non-zero Fourier coefficients and the remaining partial derivatives:

$$\frac{\partial u}{\partial y} = \frac{\left(-2b_2 - R\frac{\partial v}{\partial x}\right)}{R} \tag{3.48}$$

$$\frac{\partial v}{\partial y} = \frac{(a_0 + a_2)}{R} \tag{3.49}$$

$$\frac{\partial u}{\partial x} = \frac{(a_0 - a_2)}{R} \tag{3.50}$$

where R=htanϕ as before. The zonal (H_x) and meridional (H_y) winds are fully resolved.

There are two more factors that need to be considered when resolving H_x and H_y: the inner ring only has 4 zones, and is therefore not sufficient to fit the six coefficients for H_x and H_y. Also the partial derivatives are calculated effectively by looking at differences across the sky and it is inappropriate to apply different values from the inner ring and the surrounding outer rings. As a compromise, only the m=1 coefficients, i.e. u_0 and v_0, are evaluated independently for all zones. As for m=0 and m=2, the partial derivatives are first calculated for zones in the middle and outer rings, but the average of these two sets of results are used over the whole FOV and all zones when determining H_x and H_y.

The fitted wind field that lies perpendicular to the instrument's LOS can be represented by the following equation with the Taylor series results of H_x and H_y:

$$H^{\perp}(\theta_k, \phi) = H_x sin\theta_k - H_y cos\theta_k. \tag{3.51}$$

Finally, the meridional and zonal components of the actual horizontal wind field can be estimated by combining the observed LOS components (equation 3.36), $H^{\parallel}(\theta_k, \phi)$, with the fitted horizontal component perpendicular to the LOS, $H^{\perp}(\theta_k, \phi)$, by:

$$H_x(\theta_k, \phi) = \frac{H^{\parallel}(\theta_k, \phi)}{sin\phi} cos\theta_k + H^{\perp}(\theta_k, \phi)\, sin\theta_k \tag{3.52}$$

$$H_y(\theta_k, \phi) = \frac{H^{\parallel}(\theta_k, \phi)}{sin\phi} sin\theta_k - H^{\perp}(\theta_k, \phi)\, cos\theta_k. \tag{3.53}$$

It is cautioned that the final wind field contains all the features of the actual horizontal wind in the LOS direction, but it is possible that the perpendicular direction is not estimated correctly, with unknown distortion or missing part of the structure. Although this is not entirely satisfactory, visual interpretation of horizontal wind vectors is considerably easier than it is for LOS data (Conde and Smith [1998]).

Browning and Wexler [1968] further apply the partial derivatives in the Taylor series to determine the two-dimensional horizontal divergence, stretching deformation and shearing deformation in the wind field, shown respectively in equation 3.54. They are not analysed because the wind field itself is the main interest for the research in this thesis.

$$\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right), \quad \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right), \quad \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right) \tag{3.54}$$

The sinusoidal instrumentation baselines in the LOS wind are removed before any horizontal wind field analysis (section 3.8.6.9).

3.9.3 Test: relative sizes of the partial derivatives

This section tests the relative contributions of the different terms in the Taylor series for H_x and H_y. The tests investigate the assumptions used in $\frac{\partial v}{\partial x}$, as well as the relative sizes of the different derivatives to the uniform components in the Taylor series, u_0 and v_0.

Data from 1^{st} February, 2008 are chosen for the investigation as a example, because it is a good quality data set taken under clear-sky conditions. It is accompanied by co-

located FPI measurements for direct comparisons. Similar tests are performed on different nights with no significant difference from the results here.

Figure 3.29 shows data from the southern zone in the outer ring only, apart from the plotted u_0 and v_0 factors, where results from two adjacent zones are also shown for comparison. This is a typical night and relatively strong LOS winds (∼-200m/s, away from detector) were detected at ∼20 and 28UT.

The top two plots show the measured and fitted LOS winds from the southern, outer ring zone. The fitted wind is a reconstruction of LOS wind from the estimated $H_x(\theta_k, \phi)$ and $H_y(\theta_k, \phi)$ (equations 3.52 and 3.53). It should therefore give an indication of how well the fitted perpendicular to LOS components match with the measured LOS winds. A χ^2 test (second plot in the right column) is performed on the measured and fitted LOS wind, where low χ^2 indicates that the LOS measurements are a good representation of the fitted perpendicular components.

The vertical wind from the zenith central zone is also plotted (second plot in the left column), to give an indication of the assumed uniform vertical wind field used in the fitting process. Comparing this to the χ^2 results, it can be seen that χ^2 values are increased when there is large vertical wind at ∼22UT, though χ^2 values could be higher because of factors such as low intensity at other time stamps. A large vertical wind is typically an indication of non-uniform vertical wind field due to geomagnetic activities, and the perpendicular to LOS component may not be well fitted under the uniform vertical wind field assumption.

The six factors determined from the Taylor series (equations 3.38 and 3.39) are shown in rows 3 to 5 in both the left and right columns. The most obvious result is that u_0 and v_0, the uniform components in the Taylor series, on average dominate over the other terms at ∼100m/s. Terms involving the derivatives, $\frac{\partial v}{\partial y}$, $\frac{\partial u}{\partial x}$ and $\frac{\partial u}{\partial y}$, are relatively small, averaged at ∼20-50m/s, even though there are points in the time series when the partial derivative terms dominate.

The term involving $\frac{\partial v}{\partial x}$, which requires the assumptions in equations 3.46 and 3.47, is the smallest term with values averaged at ∼0.02m/s. As suggested by Conde and Smith [1998] the only way that a large temporal change could fail to be interpreted as a large $\frac{\partial v}{\partial x}$ variation is: if the actual $\frac{\partial v}{\partial x}$ spatial variation at the time almost exactly cancels the temporal change consistently. This situation is considered to be unlikely and indicates that any error associated with the $\frac{\partial v}{\partial x}$ estimation is negligible. It should also be noted that the periodic spikes in the $\frac{\partial v}{\partial x}x$ factor are related to the time delay when a LASER and

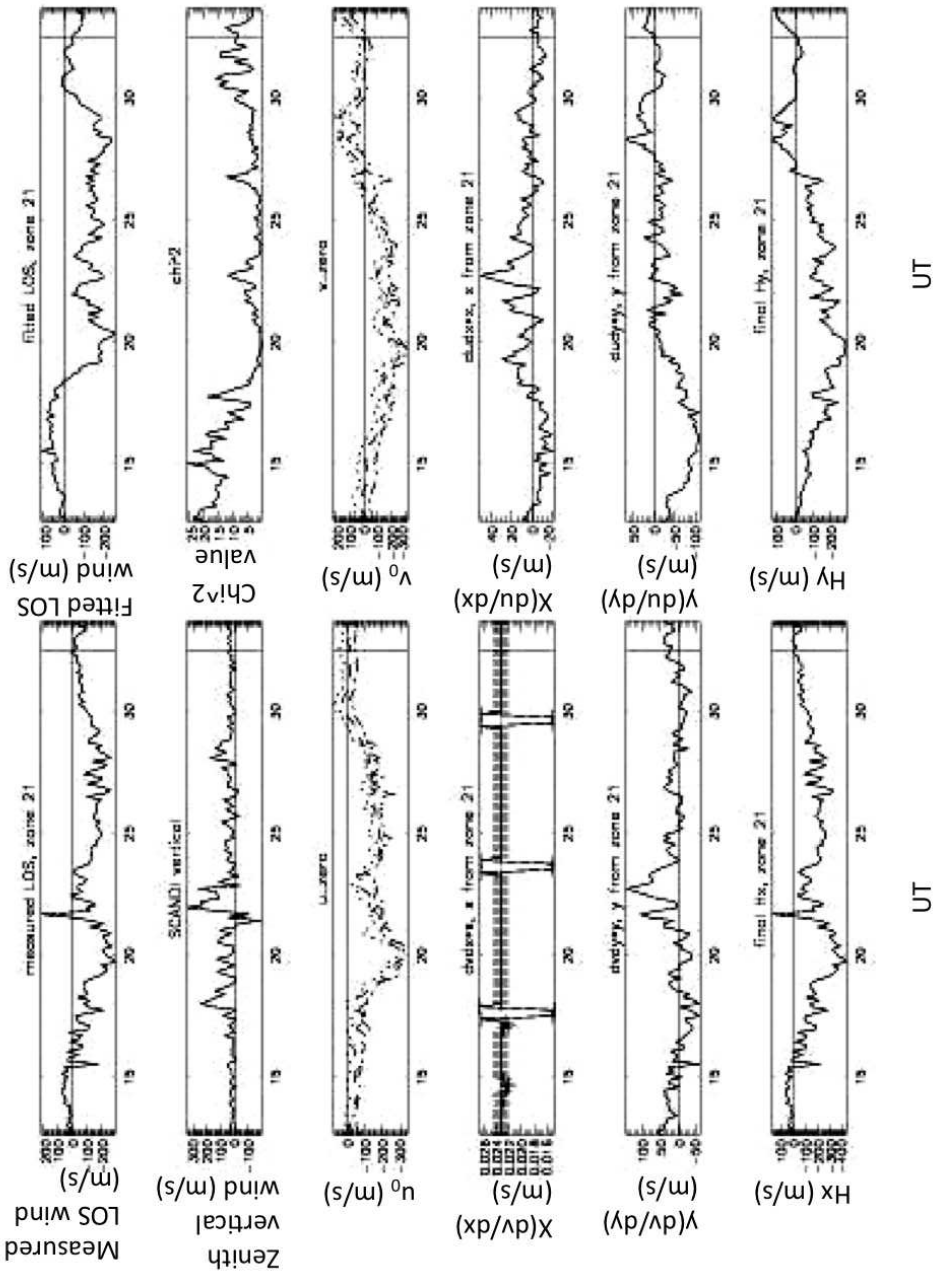


Figure 3.29: This figure tests the relative contribution of the different terms in the Taylor series described in equations 3.38 and 3.39. All results are from the southern zone in the outer ring from the night of 1^{st} February, 2008. The y-axis of each plot is in units m/s, and the x-axis are hours in UT. For the column on the left, the factors plotted from top to bottom are measured LOS wind, the zenith vertical wind, u_0, $\frac{\partial v}{\partial x}x$,$\frac{\partial v}{\partial y}y$ and H_x. The column on the right, factors plotted from top to bottom are the fitted LOS wind recalculated from the determined H_x and H_y, the χ^2 value by comparing the measured and fitted LOS wind, v_0, $\frac{\partial u}{\partial x}x$, $\frac{\partial u}{\partial y}y$ and H_y. The three graphs for the u_0 and v_0 factors includes the southern zone, as well as the two zones around it.

CAL measurements are taken between the SKY detections, which leads to a larger ∂t in equation 3.47, and are now corrected in the analysis.

Finally, H_x and H_y are plotted in the bottom two figures in both columns, and they mostly follow the general trend of u_0 and v_0 values. The test is repeated for all zones in the FOV, and gives similar results.

As a final test, simulated pure meridional and zonal wind fields are inputted into the analysis routine, and fitted LOS wind results from the outer ring are plotted against their azimuths, and are shown in figure 3.30. Both the zonal and meridional simulated inputs show the expected sinusoidal variation of LOS winds with azimuth angles for an uniform horizontal wind field, but at a different phase.

3.9.4 Test: uniform vertical wind assumption

A simple analysis is performed by comparing the χ^2 between the measured and fitted LOS wind, when the results are determined from method A: when a uniform zenith wind field is applied, determined from the zenith zone measurement; and method B: when the zenith wind is assumed to be negligible relative to the horizontal wind field and is assumed to be zero in the analysis.

The results are shown in figure 3.31. For the top panel, the measured vertical winds are plotted for the same night as the previous section for 1^{st} February, 2008. The blue diamonds mark the points in the measurements when method A (uniform vertical wind) gives better χ^2 values than method B (negligible vertical wind). Overall, method A is only better for 32% of the data points over the whole night, and are found mostly at points when the vertical winds are low. The actual χ^2 values are plotted in the bottom panel where method A results are in red, and method B results are in black. The largest increase in χ^2 are clearly found at points when zenith winds peak.

The results here indicate when high vertical winds are detected in the zenith zone, it is likely to be unrealistic to assume the same large vertical wind is seen over the whole FOV. An example could be an auroral arc appearing near the zenith causing Joule heating. Wind magnitudes can then be elevated across the zones along the zenith only.

Small zenith wind measurements give good χ^2 results as expected, but only by a small amount. It is because when zenith winds are small, they are likely to be negligible when compared to the horizontal wind field.

Hence, this gives the conclusion:

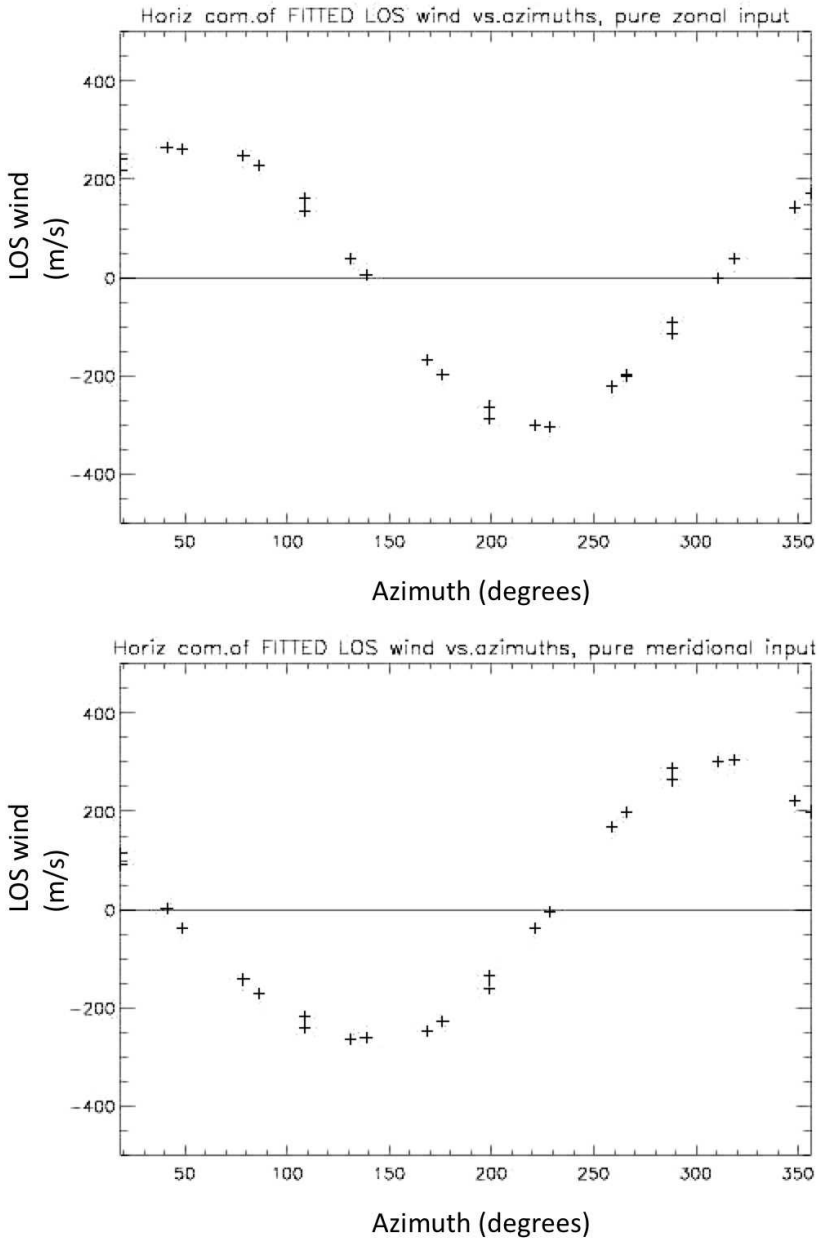


Figure 3.30: Both plots show the fitted LOS wind for zones in the outer ring, plotted against their azimuth angles. The y-axis units is m/s, and the x-axis is the azimuth angle in degrees. For the top plot, the fitted LOS wind are results from a simulated pure zonal wind field. The bottom plot is similar, but a pure meridional simulated wind field is inputted.

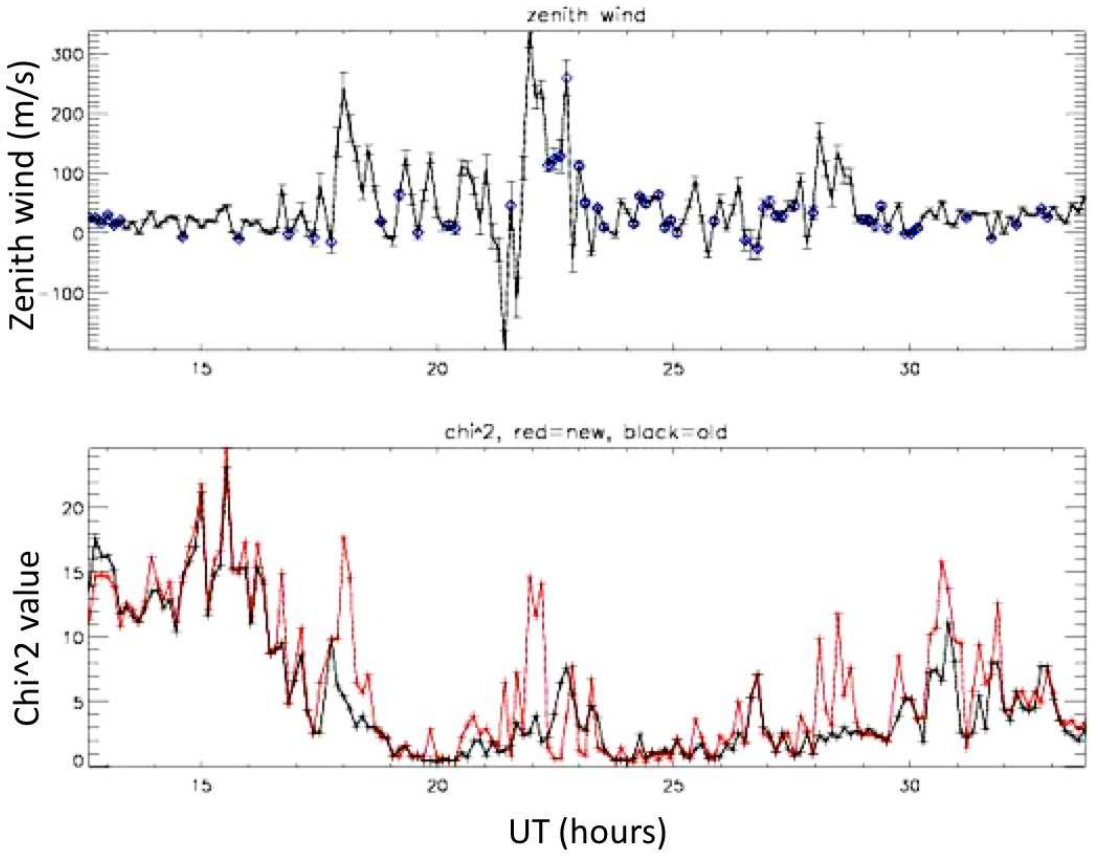


Figure 3.31: The top plot shows the measured zenith wind, with the y-axis in m/s unit, and x-axis in hours UT. Data are again from the southern outer zone from 1^{st} February, 2008. The points high lighted in blue are those when method A gives a better χ^2 i.e. an uniform vertical wind field determined from the zenith zone is assumed. The bottom plot shows the χ^2 values from the two methods, and the y-axis is represents the χ^2 values, and the x-axis remains unchanged as hours in UT. The black line represents method B, where vertical wind is assumed to be negligible, and red is method A.

- When large vertical wind is detected in the zenith zone, the estimated uniform wind field determined with the large vertical wind measurement is unlikely to be realistic.
- When the zenith winds are small, even though good χ^2 is achieved, the zenith winds are likely to be negligible in terms of the horizontal wind field anyway.
- Furthermore, only a relatively small percentage of 32% of data points are improved very slightly when method A is used.

Therefore the final decision on the analysis is that the normal analysis mode should use method B, which assumes negligible zenith wind. However, the assumptions need to

be individually tested for high geomagnetic activity data sets.

3.9.5 SCANDI-FPI comparison

Finally, the same data set from the north, south, east and west zones in the outer SCANDI ring on 1^{st} February, 2008 are compared against the co-located FPI data, both measuring red-line emission. The comparisons of the horizontal wind field are shown in figure 3.32. Please note in figure 3.32, the start time of SCANDI is earlier than that of the FPI. This allows SCANDI to be stabilised. The noisy data at the beginning of the night are collected during the warm-up period of the instrument.

The FPI and SCANDI vector results are similar. The difference contributed by the FPI narrow 1 degree FOV, compared to SCANDI's 250km zone. Also SCANDI data from the outer ring are relatively noisy due to the reduction of number of pixels within each SCANDI zone. Finally, the FPI relies only on the five LOS wind measurements at the cardinal directions plus zenith for the vector results, while SCANDI has 12 independent measurements in the outer ring. In conclusion, there is a reasonable match between FPI and SCANDI fitting methods.

3.10 Examples of SCANDI results

As a final part of this chapter, examples of SCANDI data are shown, which demonstrates different ways in which SCANDI data can be visualised. These visual displays are also developed by the author.

The simplest presentation of SCANDI results are the line plots. The neutral temperature results from the red-line data on 7th February, 2010, are shown in figure 3.33. It is difficult to see spatial variations from the line plots, but they are a good indication of the general variations of the parameters over a night. Very similar plots can be produced for the intensity and LOS wind results.

The next method is the map plot, and the spatial variations are better presented. Data from the same night in the form of map plot are shown in figure 3.34.

Finally, figure 3.35 shows an example of the horizontal vector wind field plot. In this case, for the night of 11^{th} December, 2010, SCANDI was observing red-line emission, and the co-located FPI was observing green-line emission. The two sets of vectors from the two instruments are over-plotted onto each other, which represents data measured at the F and

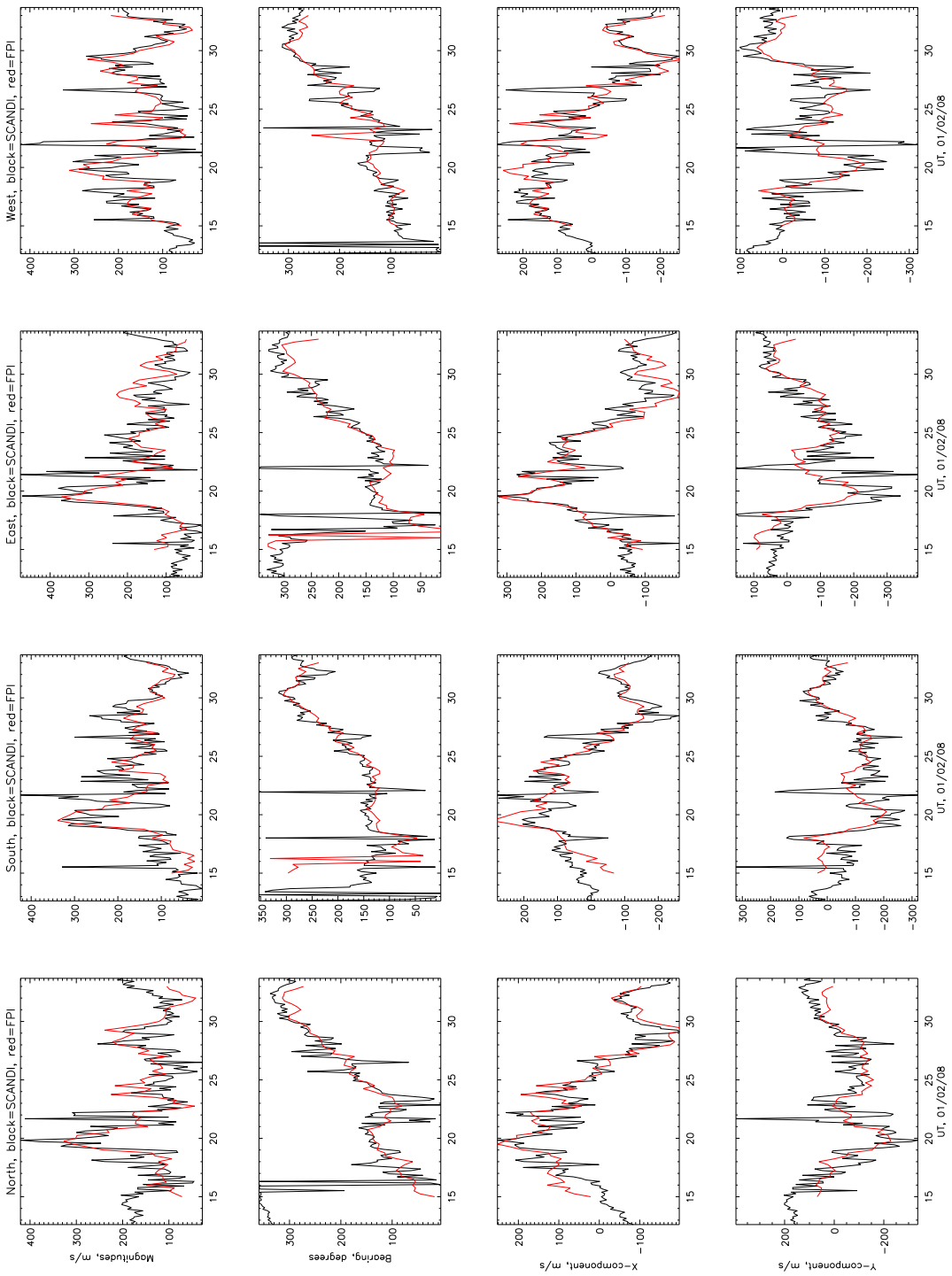


Figure 3.32: These are comparisons of the co-located FPI and SCANDI horizontal wind vectors, with FPI data in red, and SCANDI data in black. Each column, from left to right, shows the measurements from north, south, east and west directions respectively. The top to bottom rows are the magnitude of the wind vector, the vector bearing, the x component of the wind vector, and the y component. These are data from the night of 1^{st} February, 2008.

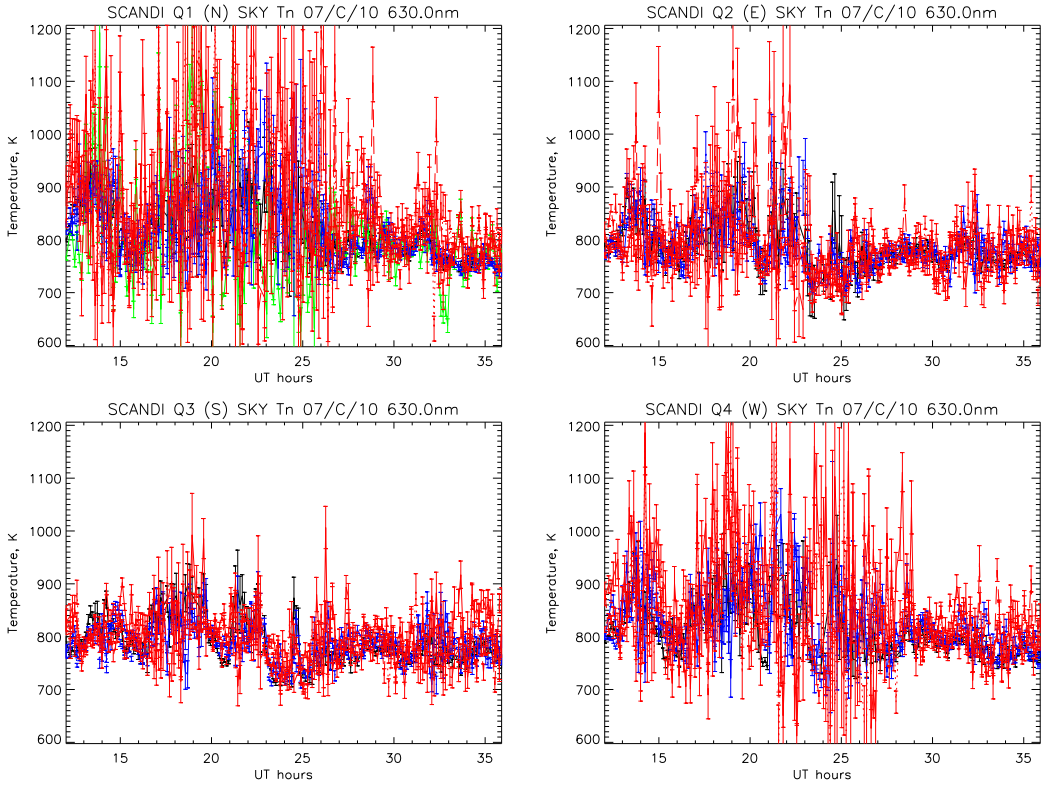


Figure 3.33: This is a typical SCANDI neutral temperature line plot, taken from the night of 7th December, 2010. Each plot shows data for zones from a SCANDI quadrant at north (top-left), east (top-right), south (bottom-left), and west (bottom-right). For each quadrant/plot, there are one zone in the inner SCANDI ring (black), two zones in the middle ring (blue) and three zones in the outer ring (red). In addition, the zenith zone is plotted in green.

E-regions for red and green-line respectively. Furthermore, a blue line is included which points towards the anti-sunward direction. This is because the F-region red-line wind field is expected to approximately follow the anti-sunward direction, since solar-pressure force dominates at most times.

It is also possible to directly compare SCANDI data to data from other instrumentation, as well as model results, as shown in figure 3.36. In this case, a typical SCANDI map plot includes the additional data of the poleward auroral oval boundary, determined from OVATION model by the Johns Hopkins University Applied Physics Laboratory. IMAGE magnetometer data from the Tromsø Geophysical Observatory of University of Tromsø are also shown. Furthermore, the FPI and SCANDI wind vectors, both from red-line emission, are over-plotted onto the intensity map.

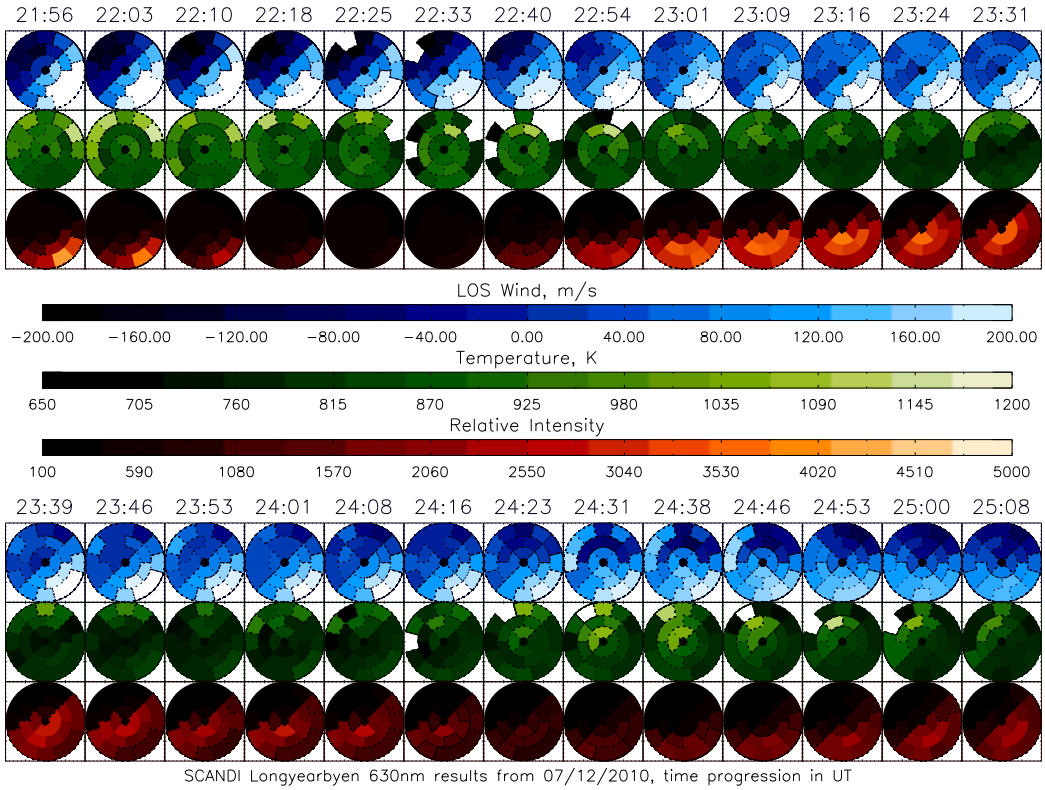


Figure 3.34: This is a typical SCANDI map plot, taken from part of the night of 7th December, 2010. Each map corresponds to SCANDI data from an exposure, where the top is north. The line-of-sight winds are plotted in blue, temperature in green and intensity in red.

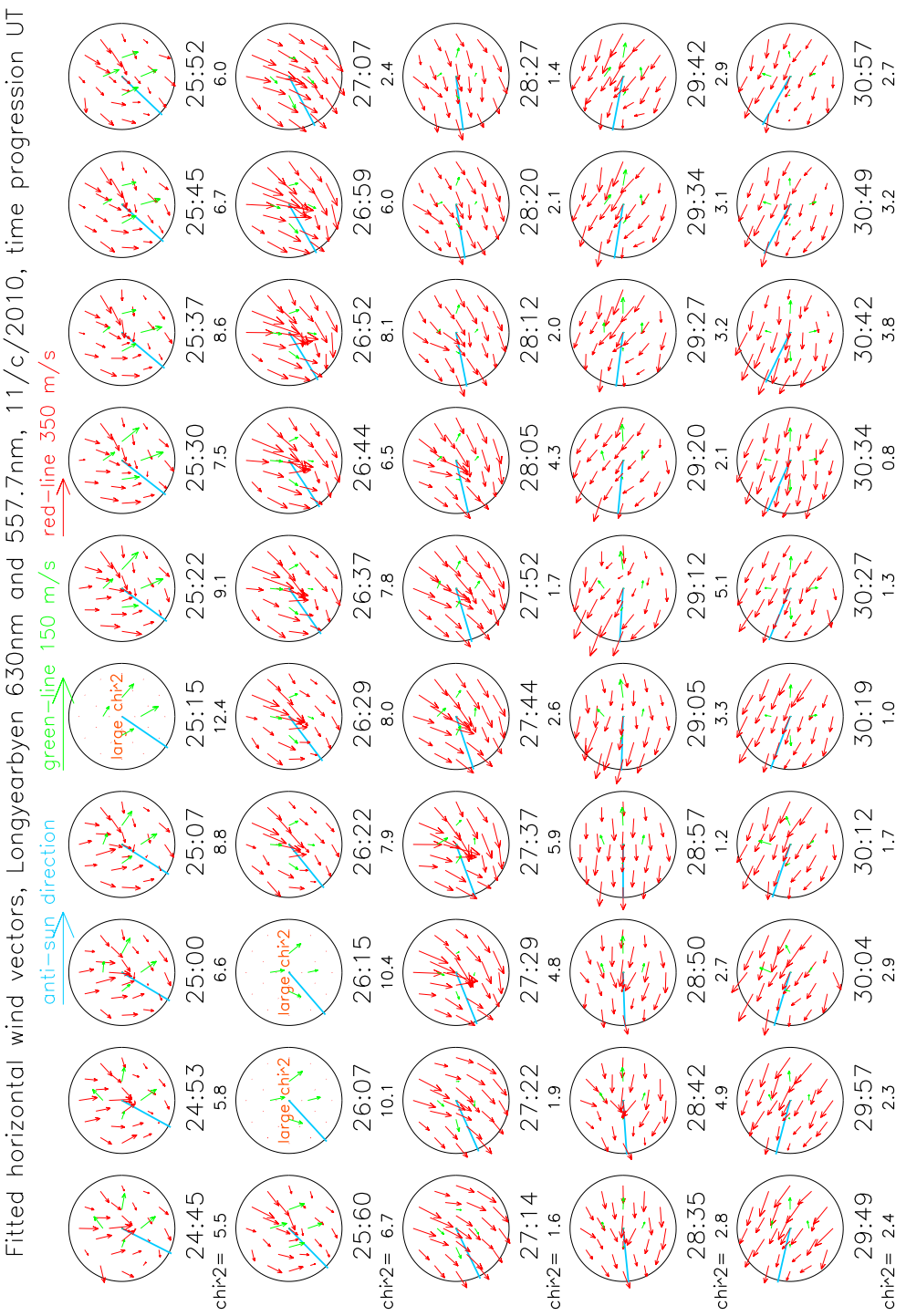


Figure 3.35: A typical SCANDI horizontal wind vector plot is shown here, taken from part of the night of 11^{th} December, 2010. Each circle corresponds to SCANDI data from an exposure, where the top is north. In this particular case, SCANDI was observing red-line emission, and its vectors are plotted in red. There were simultaneous FPI data, which were observing green-line emission available, and these are plotted in green. The blue line in each plot points towards the anti-sunward direction.

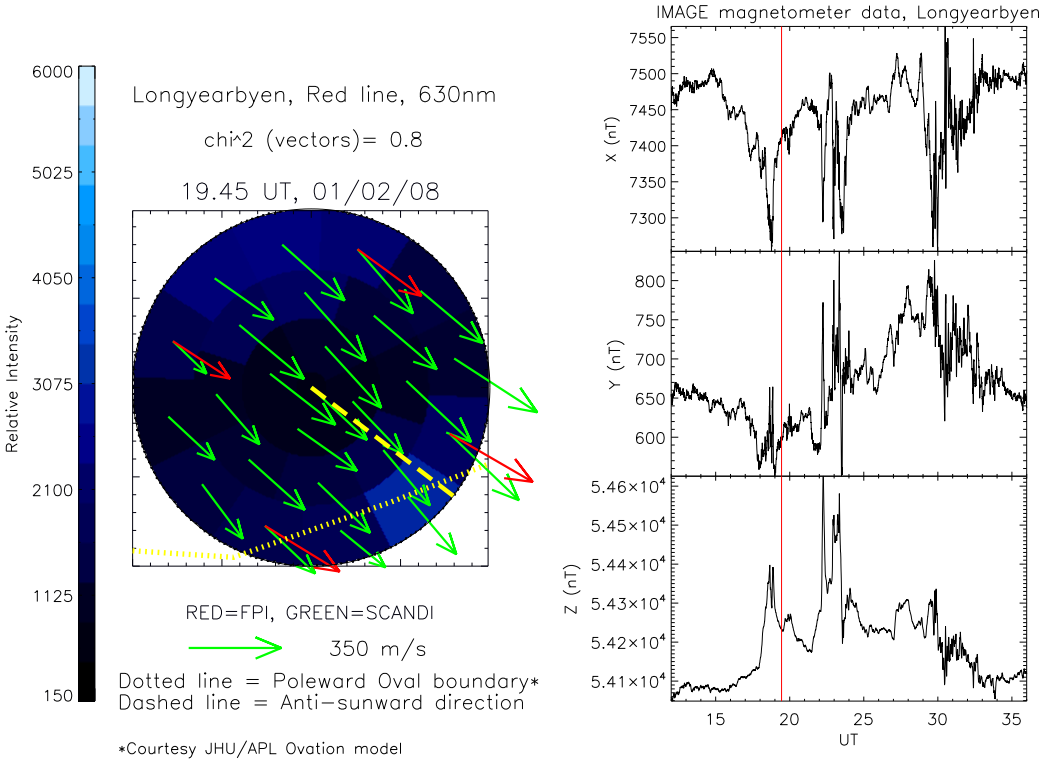


Figure 3.36: This figure shows an example of combining SCANDI data with information from other instruments. An exposure from the night of 1^{st} February, 2008 is shown. On the left, the blue background in the map shows SCANDI intensities, and top of the map is north. Both SCANDI and FPI are observing red-line emission even though they are plotted in red (FPI) and green (SCANDI). The yellow dash line points towards the anti-sunward direction, and the yellow dotted line marks the poleward auroral oval boundary deduced from the OVATION model data. On the right is the IMAGE magnetometer data taken from the same night, in the x (zonal), y (meridional) and vertical (z) direction. The red vertical line marks the time when the SCANDI exposure was taken.

Chapter 4

FPI neutral temperature trend study in the polar region

4.1 Introduction

The UCL Fabry Perot Interferometers have been measuring the red line 630.0nm emissions in the high latitude region for 30 years. They readily provide F-region neutral wind measurements from the Doppler shifts of the FPI line profile, where the temperatures can be deduced from the Full-Width-Half-Maximum (FWHM). However, determination of the absolute values requires further analysis with an instrumental calibration function, usually provided by a laser. There are many detailed studies on long-term FPI winds in the polar region available such as Aruliah et al. [1996] and Killeen et al. [1995]. There are also long-term FPI temperature investigations, for example, Smith and Hernandez [1995] and Killeen et al. [1995]. However, the temperature studies tend to involve relatively short data sets of typically half a solar cycle, while others are case studies that allow deeper analysis of wind and temperature comparisons. This chapter utilises the UCL FPI database of ten years of red line FPI neutral temperature data from years 1999 to 2008 measured at Longyearbyen, Svalbard (78N 16E). They are analysed and compared to MSIS model [Hedin, 1987, 1991] results. These years are chosen because they have readily available laser measurements for fully calibrated temperatures. This period also covers the recent extended solar minimum from cycle 23/24 at 2008, with low mass density according to Emmert et al. [2010] and low ion temperatures according to Heelis et al. [2009], resulting in an unusually cool and contracted thermosphere. Note that the Emmert et al. [2010] results are deduced from global average density data determined from satellite measurements covering years 1967-

2010. The Heelis et al. [2009] paper considers the O+/H+ transition height during June to August 2008, because the O+/H+ transition height is a sensitive indicator of solar EUV ionising flux and the topside ionosphere dynamics. Therefore, the direct neutral temperature measurements from the FPIs are a unique data set, and able to give further insight to the effect of the extended solar minimum and the temperature variation over this solar cycle.

Furthermore, the ten-year data study here could be treated as a preliminary study for longer term FPI neutral temperature trend investigations. This study is preliminary because in order to deduce an accurate temperature trend, a longer data set is required to remove effects such as the eleven-year solar cycle variation. There are numerous studies on the long term dynamics and changes of the upper thermosphere; for example, through the investigations of altitude changes of the ionospheric F2-peak from ionosonde data [Rishbeth et al., 2000, Ulich and Turunen, 1997], and trends in incoherent scatter radar measurement of ionospheric temperatures [Holt and Zhang, 2008]. A summary of different available methods for long-term trend studies can be found in Laštovička et al. [2008]. The common goal of these studies is to investigate the effect of CO_2 on the upper atmosphere, where the infrared radiative cooling by CO_2 could be a dominant effect. An increase in greenhouse gases could result in a drop in temperature in the upper thermosphere and hence a drop in height of the ionospheric F2-peak. The ten year investigation here can give an indication of whether the FPI temperature data are suitable for temperature trend analysis with the UCL FPI database of 30 years of data.

There is another application of FPI temperatures that is not discussed in this section, which is the possible effect of sudden stratospheric warming (SSW) on the thermosphere. In a SSW event, the large scale mean zonal wind suddenly slows or even reverses direction, causing a change in planetary wave activity and a rise in temperatures at stratospheric heights [e.g. Azeem et al. [2010], Goncharenko and Zhang [2008] and references therein]. SSW are mostly seen in the northern winter polar region, where it is so far inconclusive if SSW has any influence on the thermosphere [Goncharenko and Zhang, 2008]. The temperature trend study here provides the essential baseline of FPI temperatures for SSW study, where the average background variations in the FPI temperatures can be subtracted from the data set, in order to identify and isolate thermospheric SSW influences. Although not discussed in this chapter, an SSW investigation on the FPI neutral temperature was performed, with a similar method as that in Goncharenko and Zhang [2008]. However, no

SSW signature was detected. In general, it is more accepted now that SSW events can have a significant effect on the thermosphere [Goncharenko et al., 2013].

To summarise, the temperature trend chapter is divided into two parts. The first part includes a description of the analysis and preparation of the FPI temperature data; and analysis of the solar cycle and average diurnal neutral temperature variations. The analysis also presents discussions of the possible physical causes and comparisons with MSIS model results. The second part includes the average temperature results which are correlated with geomagnetic and solar indices that are frequently used in atmospheric models, with the aim to suggest the best proxy for neutral temperature in modelling work.

4.2 FPI temperature data analysis and preparation

The details of how neutral temperatures are deduced from the FWHM of the measured 630.0nm/557.7nm line shapes are described in the instrumentation chapter. The method includes the application of laser instrumentation functions to deduce the calibrated temperatures. The basic process is that the 630.0nm/557.7nm photons emitted have a Maxwellian-Boltzmann distribution of velocities and Doppler wavelengths, representing the temperature and bulk motion relative to a ground observing point (section 2.1.1). This section concentrates on the statistical analysis of the ten years of FPI temperature data, with emphasis on suitability of cloudy and moon-lit-conditions in statistical temperature studies. The effect of changes in emission profile on the FPI-measured temperature is also discussed.

4.2.1 Suitability of statistical temperature data from cloudy nights

It is well known that FPI measurements are affected by atmospheric scattering effects during cloudy conditions or from other light sources like the moonlight. Wind data are particularly sensitive [Abreu et al., 1983]. If the emission photons are above a cloudy layer, the cloud scatters and mixes the light from different directions, and hence the directional information for the bulk movement/Doppler shift information is completely lost. The general practice for wind data is to exclude all data from cloudy nights, although Abreu [1985] provides a possible inversion technique by estimating the scattering effect. For temperatures, any spatial difference is also lost in the mixing process from the cloud-scatter, but an average temperature may still be valid. Hence there is a possibility that

statistical temperatures from cloudy nights can remain useful.

Smith and Hernandez [1995] further argue that in the absence of wind at the source altitude, cloud particles are relatively cold and heavy compared with the thermospheric atomic oxygen atoms, and the effect of thermal broadening due to different cloud speeds relative to the observing point should be small compared to the thermospheric winds. Also tropospheric wind speed in the cloudy layer is much lower than the thermospheric wind speed. Hence the added broadening due to different cloud speeds is expected to be negligible.

However, in the presence of thermospheric wind, it is possible the thermospheric Doppler shift of the air parcel passing through the cloud layer is comparable to the Doppler width of the spectral line. A similar test is used, which is based on the methods in Smith and Hernandez [1995]. Their method hypothesises 300 m/s thermospheric air parcels passing through the cloud layer. It also takes the extreme case where the cloud layer scatters two wind parcels to opposite ±300 m/s velocities relative to the observing point, in additional to a stationary air parcel of 0 m/s. Note that the typical F-region wind magnitude is around 200m/s, so the estimation here represents an upper limit. The three air parcels at ±300 m/s and 0 m/s are then detected by the FPI, assumed to be coming from the same direction.

The test described here is also adjusted from that of Smith and Hernandez [1995] to values suitable for the UCL Svalbard FPI, which has a free-spectral-width of 117 bins, and each bin is equivalent to ∼45m/s of Doppler shift. Furthermore, a typical 1000K FPI spectral line shape has a FWHM of 50 bins. Hence considering the simplified case of adding the three different Doppler-shifted line shapes together, it is possible that the final spectral line would be (300/45)=6.6 bins wider on both sides. This gives a temperature that is 100*(2*6.6/50)=25% higher than the actual 1000K for this extreme case.

Smith and Hernandez [1995] then further analyse South Pole station data by comparing average results from the separated cloudy and non-cloudy data. They found that the FPI data were 1064K for the cloudy data set, and 1037K for clear-sky set, and the daily averaged clear and cloudy temperatures are indistinguishable to better than 95% confidence level. The hourly average results are used here for the UCL FPIs, as well as the daily average results, as tested by Smith and Hernandez [1995]. Hence it is necessary to test the sensitivity of the temperature data for the UCL FPIs for both sets of averages.

4.2.2 Testing cloud scattering effect on Svalbard FPI temperatures

The neutral temperature measured by the Svalbard FPI, for seasons 2007 and 2008 are tested here, where an UCL colour All-Sky Camera (ASC) was installed from 2007 onwards. The ASC provides simultaneous all-sky images at five minutes time resolution and hence the sky conditions can be deduced. The Svalbard FPI season begins at the beginning of October and ends at the end of March in the following year.

The all-sky camera data are inspected by eye, and 'cloud flags' are constructed providing the hourly sky conditions. These flags are then used to distinguish which FPI data are from cloudy or clear sky conditions, as well as data that have no ASC information.

The daily average and hourly average (separated according to universal time, UT) are determined for each month, and the months tested are 12/2007, 01/2008, 03/2008 for the 2007 season; and 09/2008, 10/2008, 11/2008, 12/2008 for the 2008 season, and are only limited by FPI data availability.

4.2.2.1 Monthly average data

The overall monthly averages are shown in table 4.1, with the associated standard deviations, σ, the sample sizes, n, and the 95% confidence intervals. The 95% confidence interval is calculated as $1.96^{*}\sigma/\sqrt{n}$, with the assumption that neutral temperatures are normally distributed.

Table 4.1: Monthly mean temperature data

	All data			Clear Sky Only		
Month	mean±95% confidence interval, K	σ, K	n	mean±95% confidence interval, K	σ, K	n
Dec 07	1200±3	353	50913	1094±4	290	16979
Jan 08	882±2	209	50290	844±4	178	8709
Mar 08	1033±6	339	14176	979±10	293	3435
Sep 08	1126±14	368	2749	1056±17	219	706
Oct 08	1072±4	343	26561	975±6	231	5515
Nov 08	960±3	272	41524	856±3	164	10344
Dec 08	868±2	118	12496	881±4	132	5419

Table 4.1 shows the monthly average data, the all-included data (i.e. cloudy data are

left in) give higher averages than the clear sky averages, apart from December 2008. The anomaly in December 2008 is likely to be because of contamination from moon-lit measurements. This indicates that the neutral winds during cloudy conditions are sufficient to cause the FWHM of the emission to widen, and artificially increase the measured temperatures. However, the differences in the averages still remain within 1σ of each other. The data set now shows a higher difference between cloudy and clear sky temperatures than Smith and Hernandez [1995]'s results. They show that their data sets match with 95% confidence level with a difference of 27 K only. The FPI temperature distribution is therefore examined further in the sections that follow. Three months of data with the largest samples are chosen for additional investigation. They are December 2007, January 2008 and November 2008.

4.2.2.2 Daily average data

The daily average data are shown in figures 4.1 and 4.2 for the two consecutive months: December 2007 and January 2008 in the middle of winter. December and January in Longyearbyen, Svalbard are under 24 hours of darkness, and hence the FPI can observe continuously without interruption. The data here are therefore not biased towards any particular UT, and are good representations of the high-latitude polar neutral F-region.

Figures 4.1 and 4.2 compare the daily means of the FPI neutral temperatures, between sampling from all the available data points, from cloudy sky conditions, and from clear sky conditions only. The data sets combine the five FPI observing directions i.e. North, East, South, West and Zenith. This is because there is no significant difference in the daily means for the individual directions.

Comparing the means calculated from all data points, and those from clear-sky conditions only, they match with each other within the 95% confidence interval for most days, apart from, for example, days 21 to 23 for January 2008 (figure 4.2). These are the days when the data are affected by moonlight. The effect of moonlight, especially for full moon, on the measured FPI emission line-shape is a high increase in the measured continuum.

For other instruments such as a photometer, the continuum needs to be accurately measured and separated from the measurement using an off-band filter, since the measurement is a simple number. For example, an off-band 644.6nm continuum wavelength is measured simultaneously with the red line 630.0nm emission by Vlasov et al. [2005]. For the FPIs, a complete interferogram/line-shape is measured, where the ratio of peak height

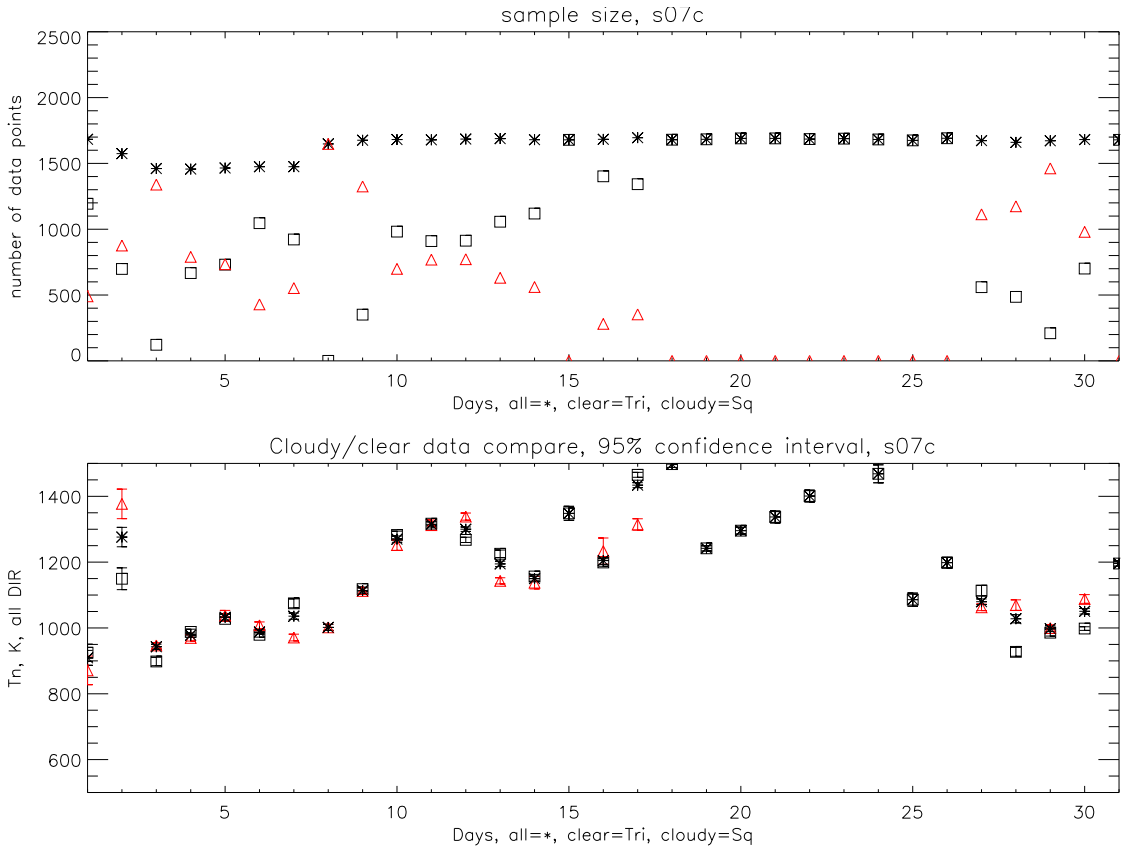


Figure 4.1: The bottom plot shows the average neutral temperatures calculated for each day in December 2007. The top figure shows the sample size. The triangles are averages/sample size from clear-sky data only, squares are from cloudy data only, and the asterisk is from all data. The clear-sky average (triangle) is also high lighted in red. The error bars are the 95% confidence interval. The cloudy effect on temperature is not as prominant here because of the general elevation of temperature around days of the full moon (24^{th} December, 2007).

to the valley height should be constant in the absence of a continuum, irrespective of signal strength. Therefore, a possible improvement to the FPI analysis routine is to take a reference peak-to-valley ratio under clear-sky conditions, and adjust the line-shape accordingly when there is moon-light. The solution is therefore to use the reference for adjustment of continuum, and match the measured line-shape with the reference peak-to-valley ratio.

At the time of the FPI temperature trend analysis, the above method has yet to be implemented, hence the moon-lit days should be excluded in the sample. Using January 2008 as an example, excluding the moon-lit data gives an overall monthly temperature mean of 848±2K , if all but the moon-lit data points are included (882±2K for data including moon-lit days in table 4.1), and a mean of 844±4K if only data from clear-sky conditions are used. Hence, the cloudy condition has a negligible effect of 4K on the data

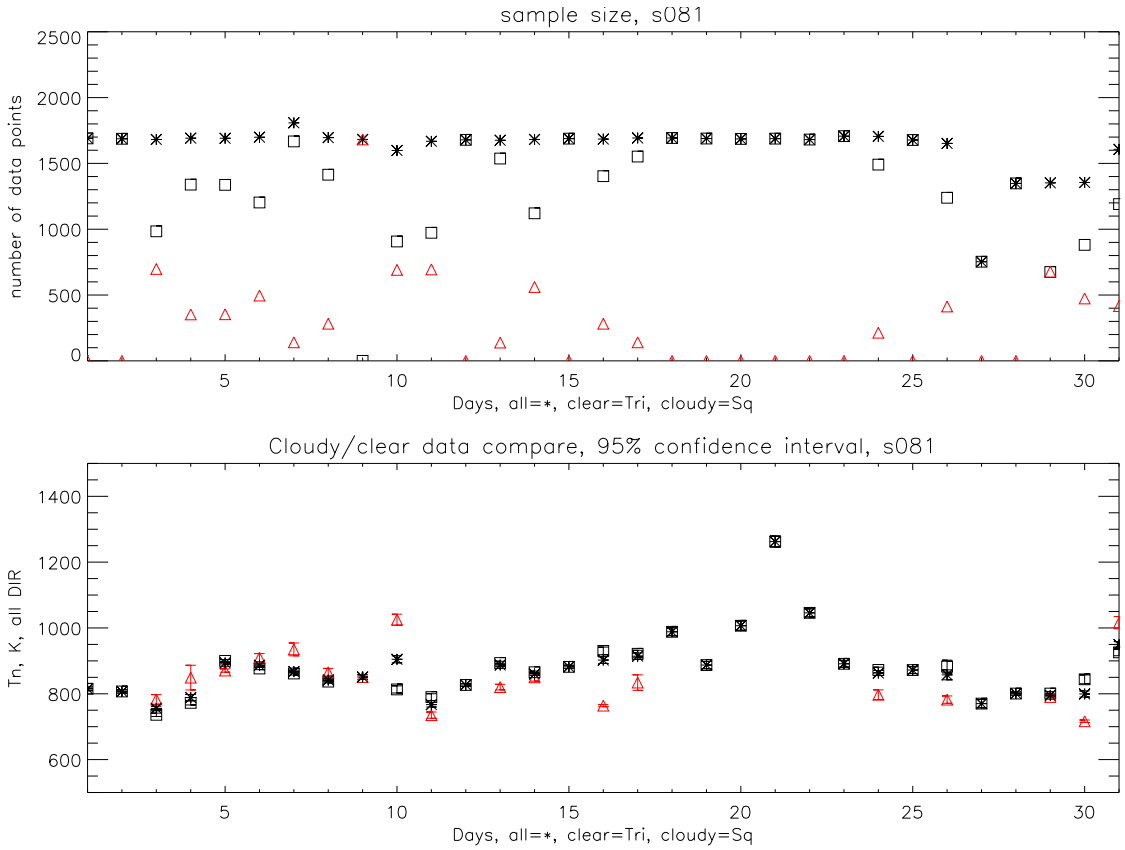


Figure 4.2: This is very similar to figure 4.1, but these are the daily means from the following month in January 2008. The full moon is at 23^{th} January 2008. Clear temperature elevation can be seen for this month because of the general lower background temperatures around the days of the full moon.

when the moon-lit data are excluded. So for a temperature trend study, if daily averages are used, it is not necessary to separate cloudy and clear-sky data, as long as the moon-lit data are excluded. The conclusion here now matches with the conclusion in Smith and Hernandez [1995], which shows that the data match with 95% confidence level.

4.2.2.3 Hourly (UT) average data

The next step is to investigate the effect of cloud-cover on the average diurnal variation of neutral temperature. The data for each month are sampled and a mean taken according to each hour of universal time (UT), and are shown in figures 4.3 and 4.4. Unlike the daily average results in section 4.2.2.2, the clear-sky sample gives a significantly lower mean temperature from the all-data-points sample. It is therefore necessary to separate the clear-sky data from the cloudy ones for a diurnal temperature statistical study.

Similar to the daily average comparisons, the temperature data from the five FPI

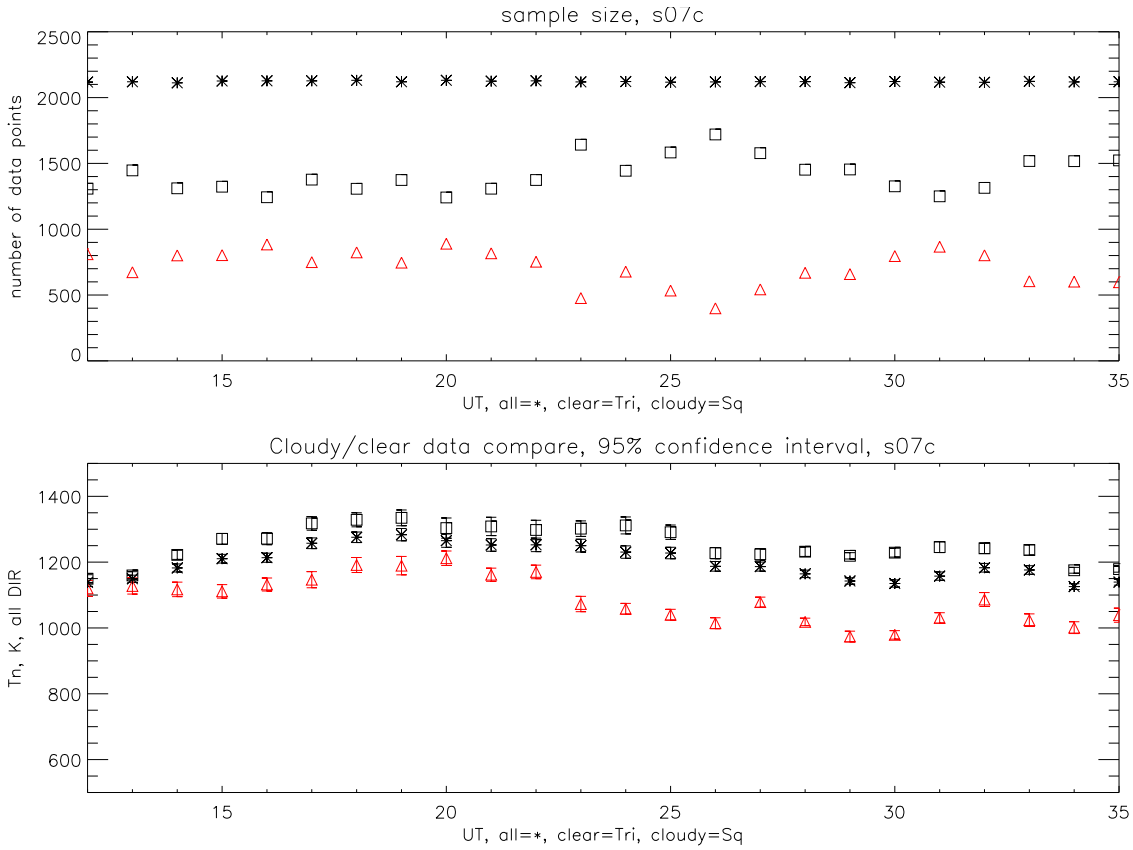


Figure 4.3: The bottom plot shows the average neutral temperatures calculated for each hour over the whole month in December 2007. The top figure shows the sample size. The triangles are averages/sample size from clear-sky data only, squares are from cloudy data only, and the asterisk is from all data. The clear-sky average (triangle) is also high lighted in red. The error bars are the 95% confidence interval.

observing directions are grouped together in figures 4.3 and 4.4, while figure 4.5 shows the results separated into the individual directions for January 2008. There is no significant difference between the average temperatures from the different directions, and so it is acceptable not to separate the directions for statistical diurnal neutral temperature studies, which is similar to the daily-average results.

The widening in the FWHM of emission line shape, and hence higher temperatures is related to the neutral wind speed. Under cloudy conditions, there should be a correlation between the magnitude of the wind and the artificially increased temperatures. Figure 4.6 shows the hourly-average diurnal neutral winds from January 2008, for the same month as the temperatures plotted in figure 4.5. Although neutral wind averages calculated from all data points, from cloudy data only and from clear sky data only are all shown in the plot, only the clear sky wind data are accurate because the winds are highly sensitive to clouds.

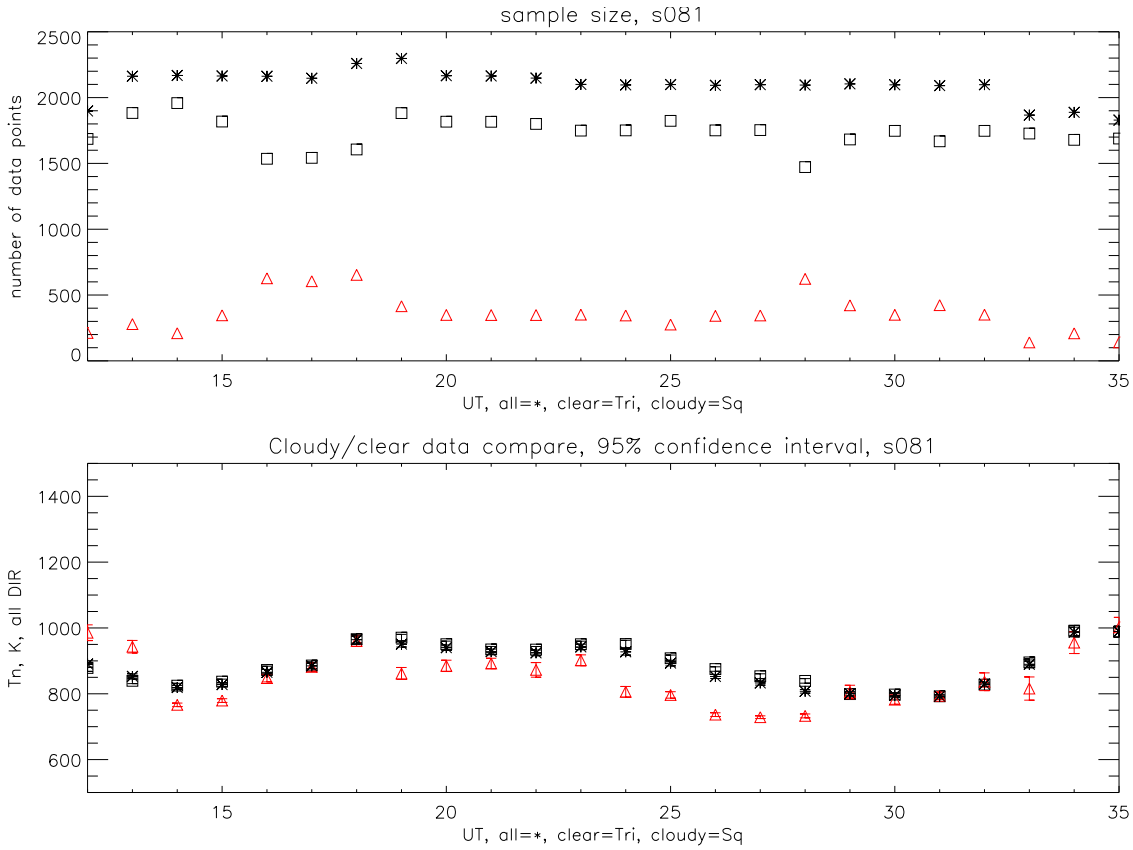


Figure 4.4: This is very similar to figure 4.3, but these are the hourly means from the following month in January 2008.

This is because clouds simply scatter the red line emission photons in all directions, and the Doppler directional information is mixed and lost. The wind information will also be lost if data from all directions are averaged together, because it is expected that the wind flowing into and out of the observing volume is equal for conservation and continuity of mass.

The winds in figure 4.6 clearly show diurnal variations for this month. Since the greater the magnitude of neutral wind, the greater the influence it has on cloudy temperatures, the cloudiness effect on temperatures will be likely also to follow a diurnal pattern. This suggestion explains why the daily averages from the previous section show negligible difference between clear-sky only data and those from all data points, because the cloudiness diurnal effect is more or less cancelled out in the daily averages. However, for hourly UT average statistical study, the diurnal wind effect on the neutral temperature under cloudy conditions will remain. Therefore it is important that cloudy data are excluded in any hourly-UT statistical temperature study.

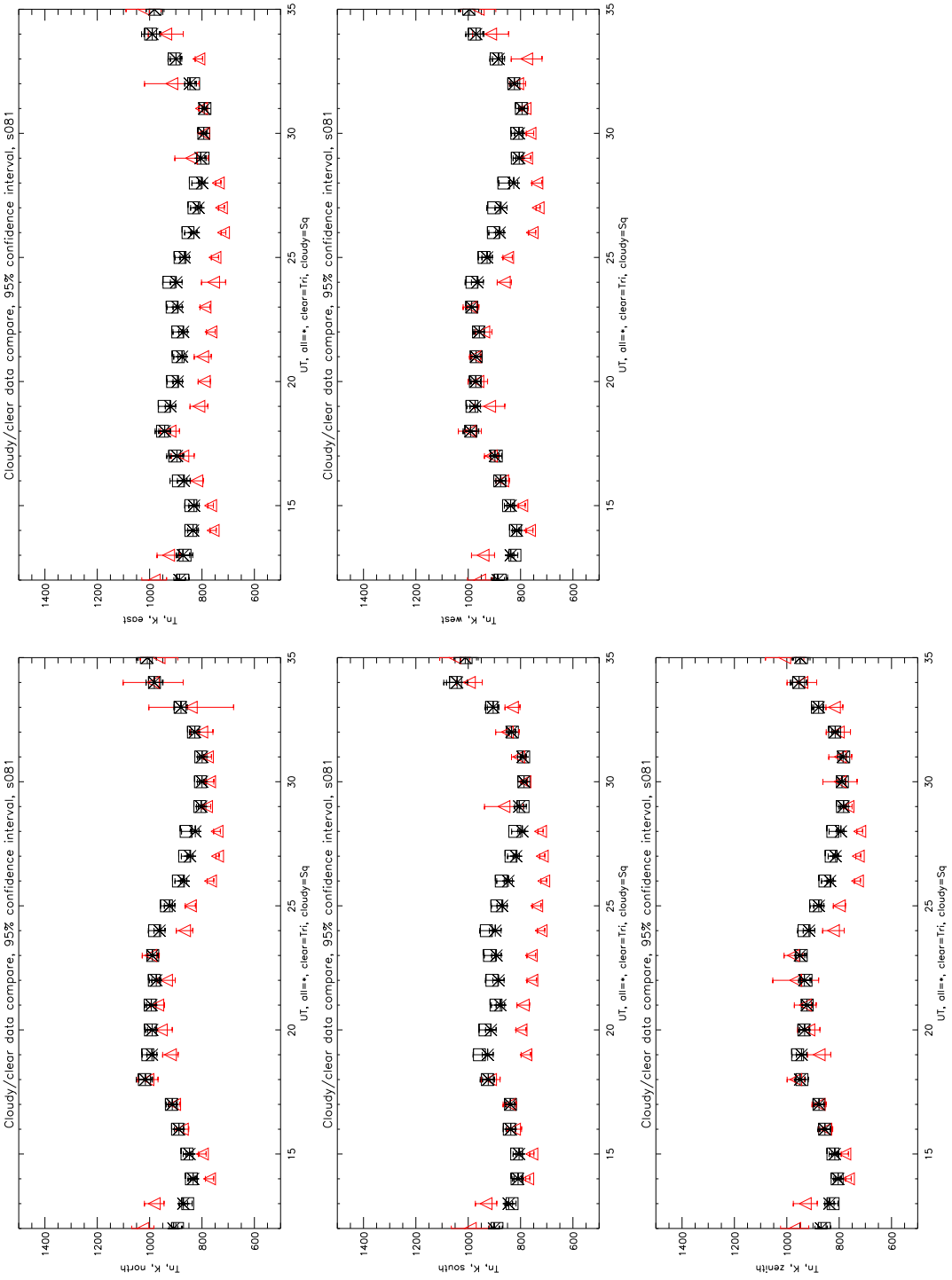


Figure 4.5: Each plot in this figure is very similar to the bottom plot in figure 4.4, where each of them shows the average neutral temperatures calculated for each hour over the month January 2008. However, the averages are separated into the five different FPI observing directions: North (top-left plot), East (top-right), South (middle-left), West (middle-right) and zenith (bottom-left).

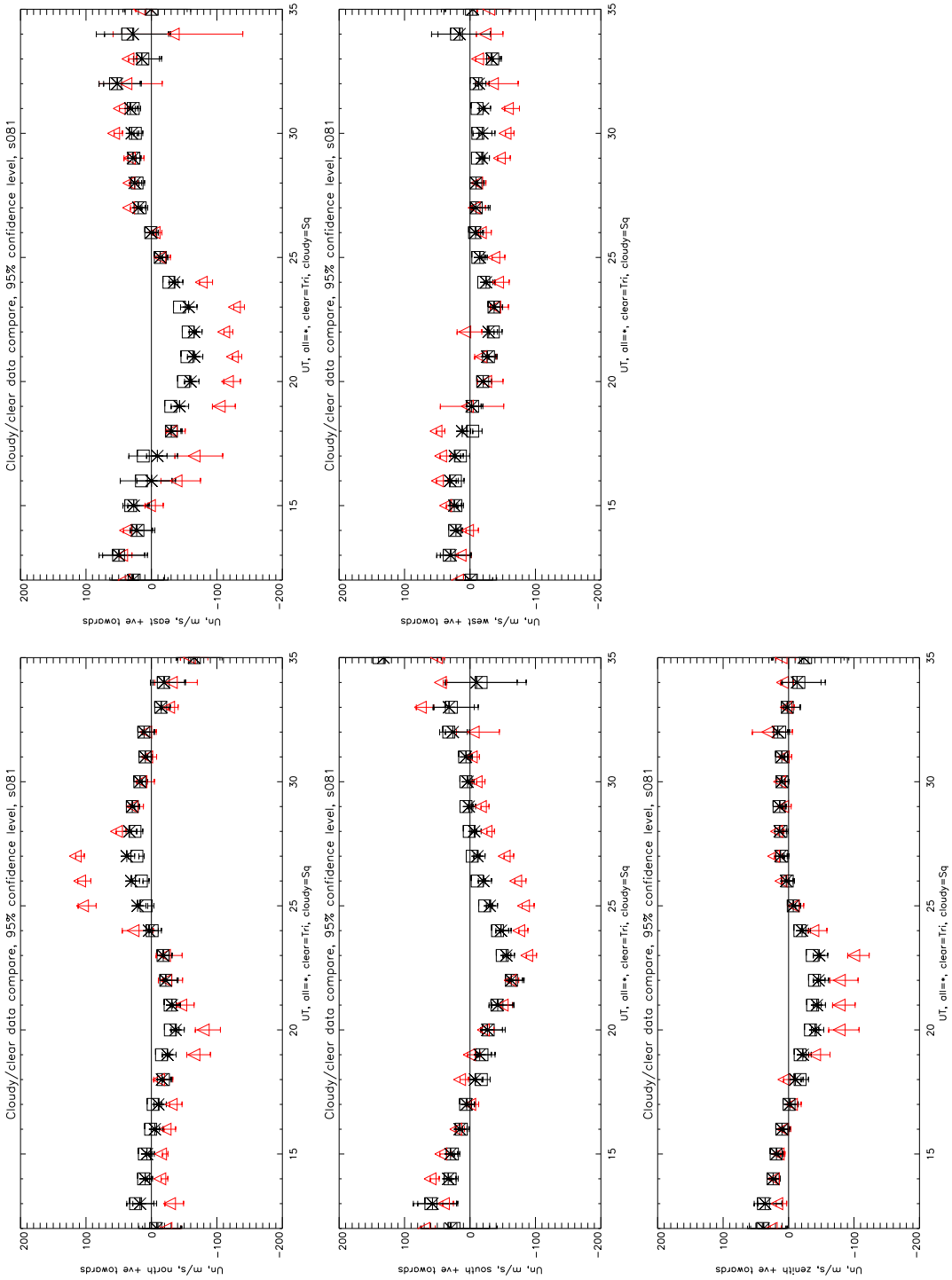


Figure 4.6: This is very similar to figure 4.5 above, but instead of temperatures, the average neutral winds are plotted for the same month. The plots have 'positive towards' convention.

In other words, in cloudy conditions, the aliasing between temperature and wind affects the spectral width in such a way that the Doppler shift measured by SCANDI, as well as the spectral width/temperature become inaccurate. However, for the temperature only,

the clouds are not expected to affect the results significantly, if the average values are sampled over a full 24-hour period.

Finally, to demonstrate the correlation of wind and temperature because of the cloud-cover effect, scatter-plots of neutral temperature (T_n) vs. wind (U_n) are shown in figures 4.7 and 4.8. Figure 4.7 shows the hourly average from clear sky data only. The hourly data are deduced by first calculating an average over each hour. Then a mean is taken over data that falls within each UT over a month. Similarly, figure 4.8 are the hourly average from all data points, inclusive of cloudy data, within a month. The clear-sky data only results are spread out without obvious correlation between the U_n and T_n. In contrast, if cloudy data are used, figure 4.8 clearly shows a much higher correlation between U_n and T_n in the UT average calculations, and is related by T_n=(817.0±0.7)+(2.55±0.01)*U_n. Hence, it can be concluded that under cloudy conditions, the FWHM and T_n are directly influenced by U_n in the UT average results.

It can also be speculated that periods of lower neutral wind are associated with lower ion-frictional heating and lower neutral temperature, whether it is cloudy or not over a day. Nevertheless, the daily-average temperature match is independently verified by Smith and Hernandez [1995] and the study here.

4.2.2.4 Summary

To summarise, in a statistical neutral temperature trend study:

- For daily average temperatures, there are no significant differences if cloudy data are excluded from the sampling or not. The daily average results agree within a 95% confidence level, provided moon-lit data are either excluded or correctly calibrated to remove the increased background level.

- This result matches with Smith and Hernandez [1995]'s South Pole FPI conclusion. Although Smith and Hernandez [1995] did not discuss the effect of moon light, and only the effect of cloud-scattered wind on the spectral FWHM is studied in their data.

- For hourly-UT average temperatures, the hourly mean neutral temperatures are significantly raised if cloudy data are included.

- In the hourly averaged data, the temperatures and winds show a linear relation, T_n = (817.0±0.7) + (2.55±0.01)*U_n, if cloudy data are included. This is caused by the

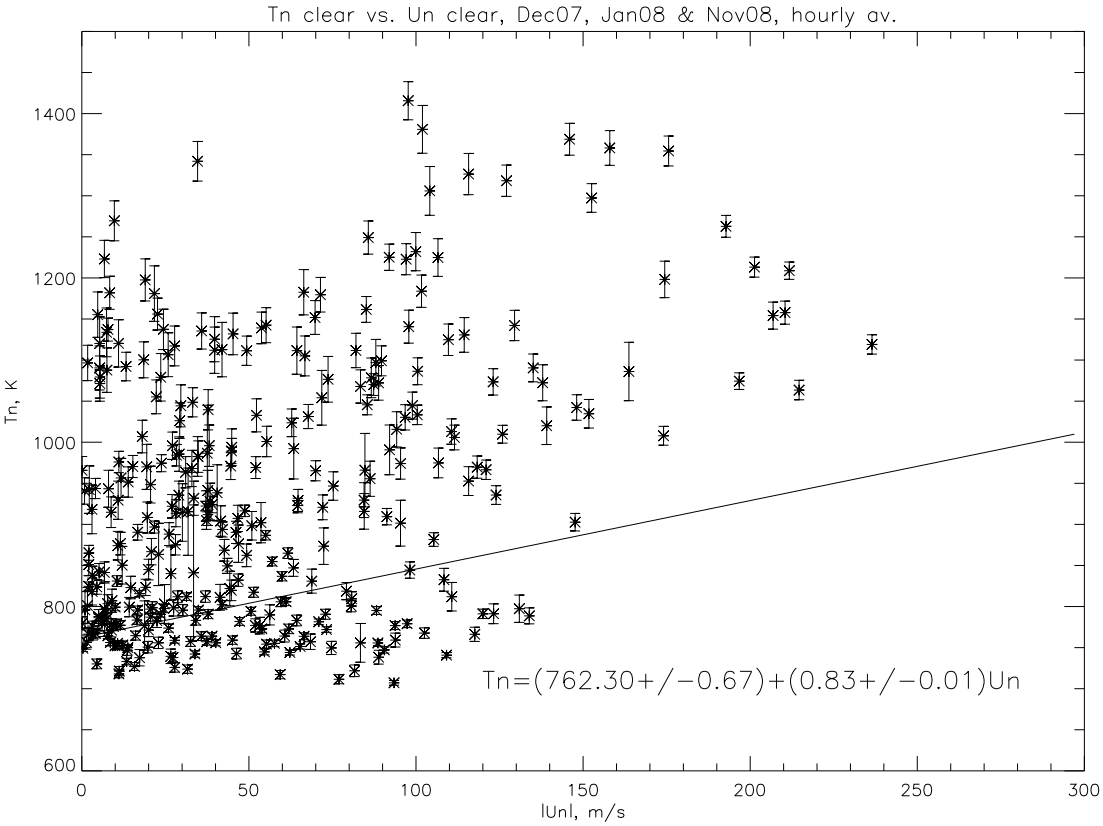


Figure 4.7: This figure shows the hourly-UT average neutral temperature vs. wind magnitude. These hourly data are deduced by first calculating an average over each hour. Then a mean is taken over data that falls within each UT over a month. These are clear-sky condition data only. Data points from each FPI direction are plotted separately. All three months, December 2007, January 2008 and November 2008 are included. The line is the first order polynomial fit and its equation.

widening of FWHM from cloud-scattered photons of significant speed.

- All directional information is lost if cloudy data are included in the analysis, for both temperature and wind.

4.2.3 Application of airglow model in FPI temperature studies

The FPI and SCANDI measure the 557.7nm (green line) and 630.0nm (red line) aurora and airglow. The neutral thermospheric winds, temperatures, intensities and background levels are determined from nonlinear least squares fits to the near-Gaussian line shape of the measured emission. The red-line profile is well understood and studied e.g. Link and Cogger [1988], Solomon et al. [1988] and Vlasov et al. [2005], and the peak emission

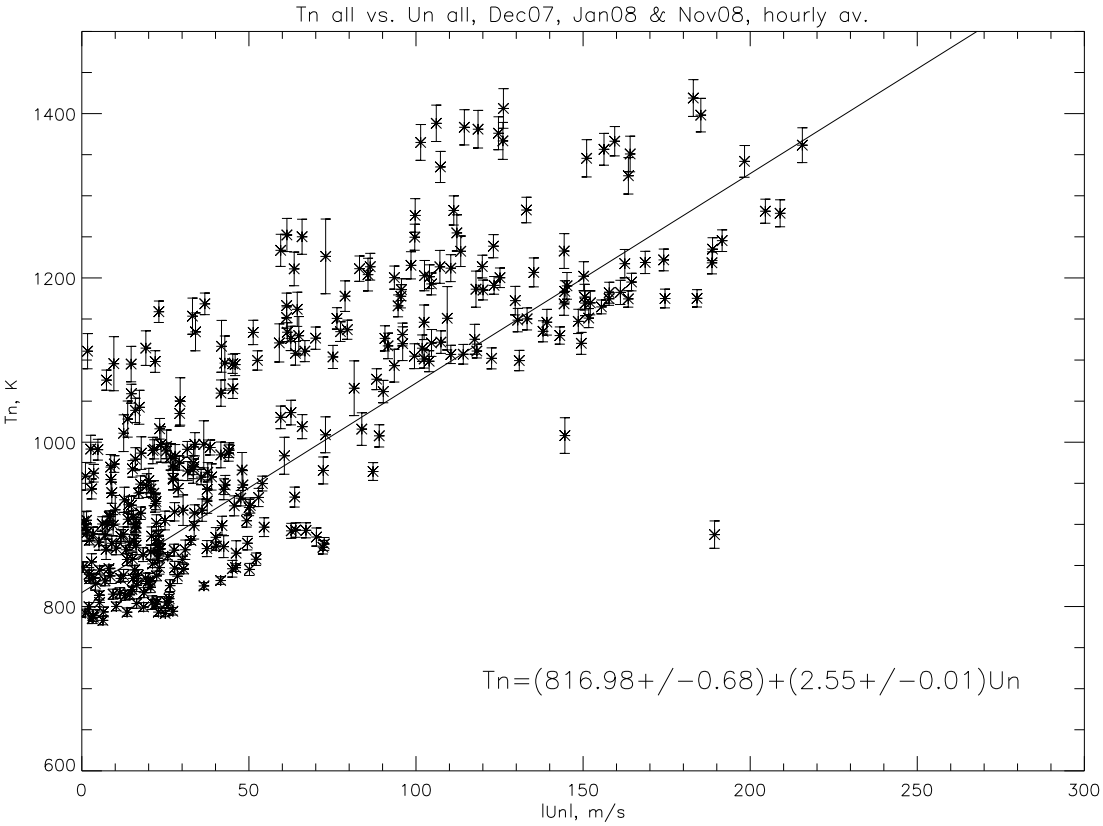


Figure 4.8: This figure is similar to figure 4.7. However, all data within a particular UT and within a month are used to determine the mean.

altitude is agreed to be at approximately 240km altitude. In addition, the average peak altitude can range from 200 to 300km for solar minimum and maximum, and has small-scale variations under different geomagnetic conditions e.g. Aruliah et al. [2005]. The green line emission is more complex, and the assumed altitude for FPI studies is about 110km altitude e.g. Kosch et al. [2010]. Other studies find the green peak emission to be lower at 95km altitude at both middle and high latitude sites e.g. Vlasov et al. [2005], Batten et al. [1988] and Schubert et al. [1999]. The shape and peak of the emission profile is important in determining the altitude at which the FPI is measuring, as well as the relative contributions from different heights to the FPI line shape. It is therefore vital in understanding what the FPI data truly represent.

FPI and SCANDI provide direct thermospheric neutral temperature measurements at the E and F regions through the red and green line atomic oxygen emissions. The long term temperature data in this chapter are taken from red line measurements. Since the emissions have a high dependence on the chemistry and densities of the thermosphere, the

630.0nm profile modelling can provide further information on the effect of the extended solar minimum of cycle 23/24. The change in emission peak altitude over a cycle is also an important factor to consider in the FPI/SCANDI neutral temperature trend study.

4.2.3.1 Instrumentation and Data

The data in this experiment come from two main instruments, the EISCAT radars which provide the ionospheric parameters, and FPI and SCANDI which provide the neutral thermosphere measurements. The location chosen for this study is Longyearbyen, Svalbard at 78^0N 16^0E in the high-latitude polar region. SCANDI has an all-sky lens of 150^0 field-of-view (FOV). Assuming a red-line emission height of about 240km, SCANDI's FOV is an approximate 1200km diameter circle, which is divided into 25, ∼150km zones. Instantaneous 'map' measurements of intensity, neutral temperatures and winds are provided by the 25 segments. For the purpose here, only data from the central zenith SCANDI zone are used.

The ESR is composed of two radar dishes, which are a 42m diameter dish with fixed field-aligned look direction, and a 32m diameter dish that is fully steerable, which can point at any azimuth at elevations above 30 degrees. The field-aligned direction of the 42m dish is at 8.4 degrees to the zenith and is therefore a good representation of the vertical measurements. It measures the height profiles of electron density, and electron and ion temperatures. The electron density is a major component in determining the emission profile, and the ion and electron temperatures are also used to produce reaction rates of the chemical processes in the emission model. Furthermore, the 42m dish data match with the central SCANDI zone, and enable simultaneous observations of the ionosphere and neutral thermosphere. Table 4.2 provides a summary of the data sets chosen to be used as yearly samples in the emission height estimation over the recent solar cycle.

4.2.3.2 Photochemical models of airglow emissions

The two main emissions that are measured by the FPI, red and green line emissions, have been reviewed and studied in detail by different methods throughout the years. For example, Witasse et al. [1999] modelled the dayglow from both red and green emissions with EISCAT data at high latitudes and modelled results are compared against Wind Imaging Interferometer (WINDII) satellite data; Vlasov et al. [2005] investigated red and green airglow in mid latitude regions, and compared the modelled results with imager

Table 4.2: Summary of Experiments

Experiments	1	2	3	4	5
Date	06 Dec 00	01 Dec 01	03 Dec 02	21 Dec 03	18 Dec 04
UT	0400-3700	0400-3700	1500-2730	0200-0600	1200-1800
Site	Svalbard	Svalbard	Svalbard	Svalbard	Svalbard
Radar mode	tau0	tau0	tau0	tau0	tau0
FPI mode	mirror sequence	mirror sequence	mirror sequence	mirror sequence	mirror sequence
FPI data	red line	red line	red line	red line	red line
Experiments	6	7	8	9	10
Date	05 Dec 05	06 Feb 07	09 Jan 08	25 Jan 09	02 Feb 10
UT	1800-3000	2030-2400	1500-4800	1700-2900	1830-2500
Site	Svalbard	Svalbard	Svalbard	Svalbard	Svalbard
Radar mode	tau0	steffe	steffe	Delta Rocket	EFOLD
FPI mode	mirror sequence	mirror sequence	all-sky map	all-sky map	all-sky map
FPI data	red line	red line	red&green line	red line	red line

detections; and Link and Cogger [1988] and Hays et al. [1978] are older papers that looked at the red line emission only. One of the studies in this thesis compares an older red line model by Link and Cogger [1988] with the emission model used by Vlasov et al. [2005], which has some different choices of the values of coefficients used. The green line model is not explicitly used in the investigations in this thesis, but it is one of the major emission lines measured by the FPI and SCANDI. The green line mode discussed by Vlasov et al. [2005] and an older model from Mullen et al. [1977] are presented in appendix E for completeness.

The red line emission comes from the transition of atomic oxygen from the O(^{1}D) state to the O(^{3}P) state. The height profile of volume emission rate can be calculated, and with the assumption of photochemical equilibrium, it can be expressed as

$$I_R = A_{630.0}[O(^1D)] = A_{630.0}\frac{\sum_{i=1}^{n} P_i^{630.0}}{\sum_{j=1}^{m} L_j^{630.0} + A(^1D)} \quad cm^{-3}s^{-1} \tag{4.1}$$

where

- I_R the volume emission rate of the red-line emissions ($cm^{-3}s^{-1}$);
- $A_{630.0}$ = the Einstein transition probability for the red-line emissions (s^{-1}), and Link and Cogger [1988] chose $A_{630.0}$= 5.15×10^{-3}, while Vlasov et al. [2005] chose $A_{630.0}$= 7.1×10^{-3}. This could cause a significant uncertainty in the results;
- $[O(^1D)]$ = the density of $O(^1D)$ at photoequilibrium (cm^{-3}), and the squared-bracket represents densities;
- $P_i^{630.0}$ = the production rate of the processes involved in $O(^1D)$ production ($cm^{-3}s^{-1}$);
- $L_i^{630.0}$ = the loss rate of the processes involved in $O(^1D)$ collisional deactivation (s^{-1}); and finally
- $A(^1D)$ = the sum of probabilities for the emissions from the $O(^1D)$ excited states in units of s^{-1}. $A(^1D)$ is 6.81×10^{-3} s^{-1} from Link and Cogger [1988], and 9.3×10^{-3} s^{-1} in Vlasov et al. [2005].

Production rates

For the night time thermosphere, the main source of the metastable $O(^1D)$ and $O(^1S)$ atoms is dissociative recombination of O_2^+:

$$O_2^+ + e_{th} \rightarrow O(^3P) + O(^1D,^1S). \tag{4.2}$$

Its production rate is expressed as

$$P_1^{630.0}(z) = \beta_1 k_1 [O_2^+(z) N_e(z)] \tag{4.3}$$

where β_1 is the efficiency/quantum yield for the production of $O(^1D)$ atoms, k_1 is the rate coefficient, $[O_2^+](z)$ is the molecular oxygen ion density at altitude, z, and N_e is the electron density. There are a number of different values available for the yield, β_1, through experiments and laboratory work, and it should be within the range of 1-1.3. The choice of the yield value is discussed in e.g. Vlasov et al. [2005] and Link and Cogger [1988]. It is another uncertainty in determining the volume emission rates.

The rate coefficient, k_1, is taken as

$$k_1 = 1.95 \times 10^{-7} (300/T_e)^{-0.7} \ \ cm^3 s^{-1} \tag{4.4}$$

by both Vlasov et al. [2005] and Link and Cogger [1988], where T_e is the electron temperature, though they apply k_1 results from experiments by Sheehan and St.-Maurice [2004] and Alge et al. [1983] respectively.

There is also a small addition to the production of $O(^1D)$ because of dissociative recombination of NO^+, and it involves the following reactions:

$$NO^+ + e \rightarrow N(^2D) + O \tag{4.5}$$

$$N(^2D) + O_2 \rightarrow NO + O(^1D) \tag{4.6}$$

The dissociative recombination of NO^+ involves its own set of productions and losses, which are expressed together in a fraction that represents the second production rate term:

$$P_2^{630.0} = \frac{\beta_2 k_{2a}[NO^+]N_e k_{2b}[O_2]}{k_{2b}[O_2] + k_{2c}[O]} \tag{4.7}$$

where

- β_2 is the yield of $N(^2D)$ in reaction 4.5 and is taken as 0.76 [Link and Cogger [1988], and references therein]
- k_{2a} is the rate coefficient of reaction 4.5 and it is expressed as $k_2 = 3.5\times10^{-7}(300/T_e)^{0.69}$ cm^3s^{-1} [Vlasov et al. [2005], and references therein]
- k_{2b} is the rate coefficient of reaction 4.6 and has a fixed value of 5.3×10^{-11} [Link and Cogger [1988], and references therein]
- k_{2c} is the rate coefficient of deactivation of $N(^2D)$ atoms by atomic oxygen, and it has the value 4.0×10^{-13} [Link and Cogger [1988], and references therein]

Only $P_1^{630.0}$ is considered by Link and Cogger [1988], but both $P_1^{630.0}$ and $P_2^{630.0}$ are included in Vlasov et al. [2005]. However, $P_2^{630.0}$ should be a relatively small contribution at about 5% [Witasse et al., 1999].

Another small contribution to the production rate is the $O(^1D)$ atoms from the cascading of $O(^1S)$ state, since each 557.7nm photon emitted from $O(^1S)$ state leads to a $O(^1D)$:

$$O(^1S) \rightarrow O(^1D) + h\nu(557.7nm), \tag{4.8}$$

and its production rate is

$$P_3^{630.0} = A_{557.7}[O(^1S)]. \tag{4.9}$$

$P_3^{630.0}$ is only considered by Witasse et al. [1999], and has a very small contribution of ∼3% in the study. Both Link and Cogger [1988] and Vlasov et al. [2005] produced models for mid-latitude sites, while for the high latitude site experiments in this section, production from excitation of the ground state atomic oxygen from collisions with thermal electrons of sufficient energy also has to be considered. The reaction and production rates are

$$O + e_{th} \rightarrow O(^1D) + e_{th} \tag{4.10}$$

$$P_4^{630.0} = k_4[O](z)N_e(z) \tag{4.11}$$

where k_4 is the rate coefficient for this reaction, and has dependence on the electron temperature, $T_e(z)$, as well as the thermal electron density, $N_e(z)$, and is expressed as [Witasse et al. [1999] and references therein]:

$$k_4 = 0.596T_e^{1/2}\frac{9329 + T_e}{(51183 + T_e)^3}exp(\frac{-22756}{T_e}). \tag{4.12}$$

Note that this equation is deduced from the assumption that the reaction cross section of $O(^1D)$ is an exact Maxwellian distribution with energy. However, heating experiments show that this might not be true for the sub-thermal energy range between 2 and 4eV e.g. Gustavsson et al. [2002], Leyser et al. [2000]. The higher than Maxwellian-predicted cross section at these energies is caused by electron excitation of N_2, and the effect becomes more important at lower altitudes when N_2 density increases. Hence the equation here may underestimate the red line production rate.

The effect of photoelectron impact and photodissociation are not included here because the FPI can only measure in nighttime, hence only nightglow production terms are needed in this study.

Loss rates

The loss rates considered here are all associated with the collisional deactivation/quenching of $O(^1D)$ atoms which has a radiative lifetime of ∼110s, although a higher lifetime of 140 s has been shown by experiments by Solomon et al. [1988]. The reaction is described by

$$O(^1D) + X \rightarrow O + X \tag{4.13}$$

where X represents the collision species N_2, O_2, O or e_{th}, and their loss rates are

$$L_1^{630.0} = k_{N_2} N_2(z) \tag{4.14}$$

$$L_2^{630.0} = k_{O_2} O_2(z) \tag{4.15}$$

$$L_3^{630.0} = k_O O(z) \tag{4.16}$$

$$L_4^{630.0} = k_{e_{th}} N_e(z) \tag{4.17}$$

and the rate coefficients are

$$k_{N_2} = 2.0 \times 10^{-11} exp(k'_{N_2}/T_n) \tag{4.18}$$

where k'_{N_2} is 107.8 in Vlasov et al. [2005] and Witasse et al. [1999], and 111.8 in Link and Cogger [1988]. T_n is the neutral temperature.

$$k_{O_2} = 2.9 \times 10^{-11} exp(67.5/T_n) \tag{4.19}$$

in all three papers.

$$k_O = (3.730 + 1.1965 \times 10^{-1} T_n^{0.5} - 6.5898 \times 10^{-4} T_n) \times 10^{-12} \tag{4.20}$$

which is only used in Vlasov et al. [2005] and Witasse et al. [1999] with the same value.

Finally,

$$k_{e_{th}} = 1.6 \times 10^{-12} T_e^{0.81} \tag{4.21}$$

in Link and Cogger [1988], and

$$k_{e_{th}} = 6.6 \times 10^{-10} \tag{4.22}$$

in Witasse et al. [1999].

The quenching from thermal electrons is usually negligible.

4.2.3.3 Input parameters

The data that can be input into the photochemical model are a combination of modelled and measured results. The height profiles of ionospheric conditions can be provided directly by the EISCAT radar, and the parameters of interest are:

- N_e, which is used in the determination of the densities of ion species $[O^+]$, $[O_2^+]$ and $[NO^+]$ with the appropriate reaction rates and assuming charge-neutrality in the thermosphere. This method is applied in section 4.2.3.4.
- T_i in determining the associated reaction rates.
- T_e also in determining the associated reaction rates.

The height-profile of neutral parameters can be provided by the MSIS model [Hedin, 1987, 1991], because the FPI/SCANDI can only supply neutral temperatures that mainly represent the altitude where the peak emission lies. The MSIS neutral temperature profiles are adjusted according to the FPI measured temperatures at 240 km. The height profiles of parameters that can be used in the emission model include:

- Neutral densities of $[O_2]$, $[N_2]$ and $[O]$
- Neutral temperatures, T_n

4.2.3.4 Determining the ion compositions from the electron density profile

This is the method used by both Link and Cogger [1988] and Vlasov et al. [2005], and the ion density calculations here are adapted from their methods, which applies the measured electron density profile directly by assuming the ionosphere is charge-neutral at all times, and hence

$$[e] = [O^+] + [O_2^+] + [NO^+]. \tag{4.23}$$

The molecular ion densities are determined through the consideration of the charge exchange reactions, and their respective rate coefficients, $k"$:

$$O^+ + O_2 \rightarrow O_2^+ + O \tag{4.24}$$

$$k_1" = 3.23 \times 10^{-12} exp\left(3.72(\frac{300}{T_i}) - 1.87(\frac{300}{T_i})^2\right) \tag{4.25}$$

for Link and Cogger [1988], but Vlasov et al. [2005] chose a different $k_1"$:

$$\begin{aligned} k_1" &= 2.82 \times 10^{-11} - 7.74 \times 10^{-12}(T_{eff}/300) \\ &+ 1.073 \times 10^{-12}(T_{eff}/300)^2 - 5.17 \times 10^{-14}(T_{eff}/300)^3 \\ &+ 9.65 \times 10^{-16}(T_eff/300)^4 \end{aligned} \tag{4.26}$$

$$T_{eff} = 0.667T_i + 0.333T_n. \tag{4.27}$$

The other dissociative recombination reaction is:

$$O^+ + N_2 \rightarrow NO^+ + N. \tag{4.28}$$

$$k_2^{"} = 2.78 \times 10^{-13} exp\left(2.07(\frac{300}{T_i}) - 0.61(\frac{300}{T_i})^2\right). \tag{4.29}$$

The dissociative recombination reactions, and their respective rate coefficients, α, are:

$$O_2^+ + e \rightarrow O + O \tag{4.30}$$

$$\alpha_1 = 1.95 \times 10^{-7}\left(\frac{300}{T_e}\right)^{-0.7} \tag{4.31}$$

and

$$NO^+ + e \rightarrow N + O \tag{4.32}$$

$$\alpha_2 = 4.00 \times 10^{-7}\left(\frac{300}{T_e}\right)^{-0.9}. \tag{4.33}$$

The rearranged production equations can be written as:

$$[O_2^+] = \frac{k_1^{"}[O^+][O_2]}{\alpha_1[e]} \tag{4.34}$$

$$[NO^+] = \frac{k_2^{"}[O^+][N_2]}{\alpha_2[e]}. \tag{4.35}$$

Finally, by substituting equations 4.34 and 4.35 into equation 4.23, it gives

$$[e] = [O^+]\left[1 + \frac{k_1^{"}[O_2]}{\alpha_1[e]} + \frac{k_2^{"}[N_2]}{\alpha_2[e]}\right]. \tag{4.36}$$

The radar measured T_e and T_i profiles are also used in the reaction rates involved in the ion composition determination.

All the parameters in equation 4.36, including [e], can be deduced with either MSIS or EISCAT data, apart from $[O^+]$. Hence rearranging the equation gives the value of $[O^+]$ and all three densities of the major ionospheric species can be calculated.

4.2.3.5 Sources of uncertainties

The temperatures and neutral densities provided by the MSIS model are the first source of uncertainty. In order to minimise the temperature discrepancies, the MSIS T_n result for the altitude closest to the FPI/SCANDI measured altitude can be compared against the FPI-measured T_n. The MSIS neutral temperature profile is then multiplied by a constant

temperature shift according to the modelled vs. measured T_n difference. The shape of the MSIS profile therefore remains the same. Please note the uncertainties that may be introduced by the shift of MSIS temperature profile are unlikely to be a major factor that affects the maximum emission height value, or the application of emission profile as a weighting function. As a quality check, the modelled height-integrated intensity from the method here is compared against MSP data.

The most influential factor for this method of modelling the peak of the emission is the electron density profile. The electron density shapes the height-variation of the production rate. The MSIS neutral temperature contributes to the loss rate and only appears in the rate coefficients of N_2, O and O_2 quenching (equations 4.18, 4.20 and 4.19).

A simple check is to calculate how the combined loss rates from N_2, O and O_2 may be affected, if the neutral temperature is changed by 100K from 250K to 350K. Assuming O density is $2\text{x}10^9\text{cm}^{-3}$, O_2 density is $3\text{x}10^7\text{cm}^{-3}$ and N_2 density is $6\text{x}10^8\text{cm}^{-3}$ for an altitude of 250km, the overall loss rate is only increased by 1% when the neutral temperature is increased from 250 to 350K. Hence, it is unlikely uncertainties in the MSIS temperature affect the overall result.

In terms of the MSIS densities, Witasse et al. [1999] include an adjustment factor to the MSIS-modelled neutral densities by fitting the EISCAT measured parameters to the ionospheric model, TRANSCAR. They state that the correction factors are small for quiet periods, and an adjustment of $[O]/[O]_{MSIS}$=0.90, $[O_2]/[O_2]_{MSIS} = [N_2]/[N_2]_{MSIS}$=1.20 are used on their day of worst-fitted data. The MSIS densities that are used in this thesis are not adjusted. This is because a height-independent ratio adjustment of MSIS neutral density is not expected to affect the altitude profile of the volume emission rates. Hence, it also is unlikely have any effect on the determination of maximum emission height, or the application of emission profile as a weighting function.

The airglow models described here also only consider the effect of thermal electrons, and are not suitable for the description of the red line emission when energetic electron precipitation is present. Under aurora, the main error in applying the airglow model is that the whole of the EISCAT-measured electron density profile would be treated as if all electrons are in the thermal energy range. The reaction rates and cross sections for electrons of different energies are clearly different. Therefore additional production terms and the distribution of electron energies need to be considered. The airglow model is unable to represent this.

However, when the EISCAT-measured electron profile is applied in the airglow model analysis, the presence of energetic electrons at lower altitudes is still able to decrease the modelled peak emission altitude. The airglow model can cautiously be used as a very rough estimation of emission height variation, or simply as 'flags' for drops in emission height in FPI studies. Furthermore, periods of moderate or active geomagnetic activities (Kp $>$ 3) only account for about 10% of the time [Foster et al., 1986]. The airglow model is sufficient for the application here i.e. to estimate the general average solar cycle variation of the red line emission profile and its effect on FPI-measured temperature.

4.2.3.6 630.0nm modelled and measurement comparison

The modelled volume emission rates are compared with the Meridian Scanning Photometer (MSP) measurements, using data from experiment 8 in table 4.2. For a direct comparison, the volume emission rates are integrated for altitudes with significant red line emission:

$$J_R = \int_{z_1}^{z_2} I_R dz. \tag{4.37}$$

The altitude chosen is z_1=100km to z_2=400km, and the integration is performed with the INT_TABULATED routine in IDL, which integrates a tabulated set of data on a closed interval using a five-point Newton-Cotes integration formula [IDL, 2006]. The integrated volume emission rates, i.e. red line intensity with units Rayleigh, R, are compared against the 630.0nm MSP data (*Courtesy Geophysical Institute, University of Alaska Fairbanks*) in figure 4.9.

For the first row of figure 4.9, the plot on the left is the intensity ratio between the height-integrated volume emission rates from the two models, [Link and Cogger, 1988]/ [Vlasov et al., 2005]. The plot on the right is the intensity of the two models over-plotted in the same plot. For the second row, the plot on the left is the intensity ratio between the modelled intensity and the MSP intensity, [modelled/MSP], the red symbols are the Link and Cogger [1988] results, and black are Vlasov et al. [2005]. The plot on the right is MSP calibrated measured intensity. In the bottom row, the plot on the left is the intensity ratio between the modelled intensity and the SCANDI intensity, [modelled/SCANDI], the red symbols are the Link and Cogger [1988] results, and black are Vlasov et al. [2005]. The plot on the right is the SCANDI, uncalibrated measured intensity.

The ratios of height-integrated intensities between the two models, [Link and Cogger, 1988]/[Vlasov et al., 2005], give an average value of $\sim$0.7. There is no obvious correlation

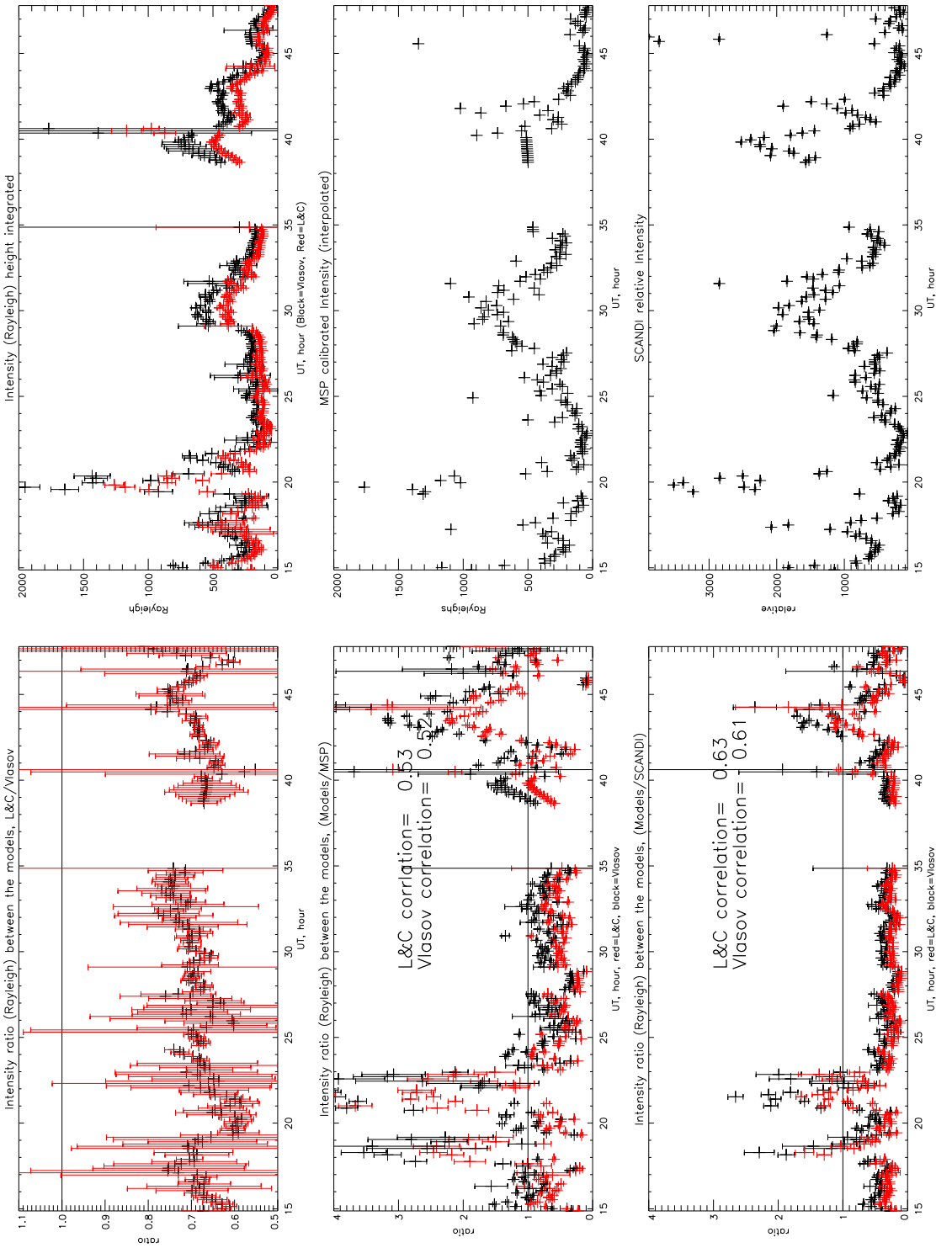


Figure 4.9: This figure summarises the height-integrated intensity results between the two models, and the measured MSP data. The left column shows the modelled/measurement ratios, and the right column shows the actual intensity values from models and measurements.

between the intensity magnitude and the difference between the two models, and the modelled-intensity results from the two models tend to follow each other with a relatively constant difference between them.

The height-integrated modelled results from both authors' methods give intensity values that are similar to the measured and calibrated MSP data. The average ratios of (modelled/measurement) in figure 4.9 is 0.94 and 1.38 for Link and Cogger [1988] and Vlasov et al. [2005] respectively. The trends are similar for both modelled comparisons with SCANDI and MSP. Note that the MSP intensity data are calibrated, but the SCANDI intensity are non-calibrated relative intensities, but the calculated ratios follow a similar trend.

Finally, the intensity correlation for the Link and Cogger [1988] model with MSP and SCANDI data are 0.53 and 0.63 respectively, and correlation coefficient for Vlasov et al. [2005] model with MSP and SCANDI are 0.52 and 0.61 respectively. The correlation shows that that modelled data are a reasonable match to measurements, considering the uncertainties listed in section 4.2.3.5.

4.2.3.7 The altitude of the red line emission peak

For SCANDI and FPI observations, the altitude of the peak of volume emission rate is an important parameter. This is because it gives the actual height of instrument observations, instead of the assumed altitude of 240km for the red line. The peak emission height from the different models, as well as a comparison with the hmF2 deduced from the EISCAT electron density data are plotted in figure 4.10. It shows that the peak altitude is very similar between Link and Cogger [1988] and Vlasov et al. [2005].

As another test of the emission height, the hmF2 altitude estimated from the electron density peak is reduced by one scale height, $k_B(T_e+T_i)/m_ig$, where k_B is the Boltzmann constant and g the gravitational acceleration. This is an approximate test for where the red line peak could be. The hmF2-estimated peak altitude falls within half a scale height ($\sim$15km) of the red line peak altitudes that are deduced from the photochemical models, for the majority of the experiment period.

The average altitude of the red line emission peak from the models is found to be about 210km, with a minimum altitude of around 190km, excluding the auroral precipitation period. The results here clearly demonstrate that the average peak is lower than the usual 240km assumed for the FPI measurements for this solar minimum.

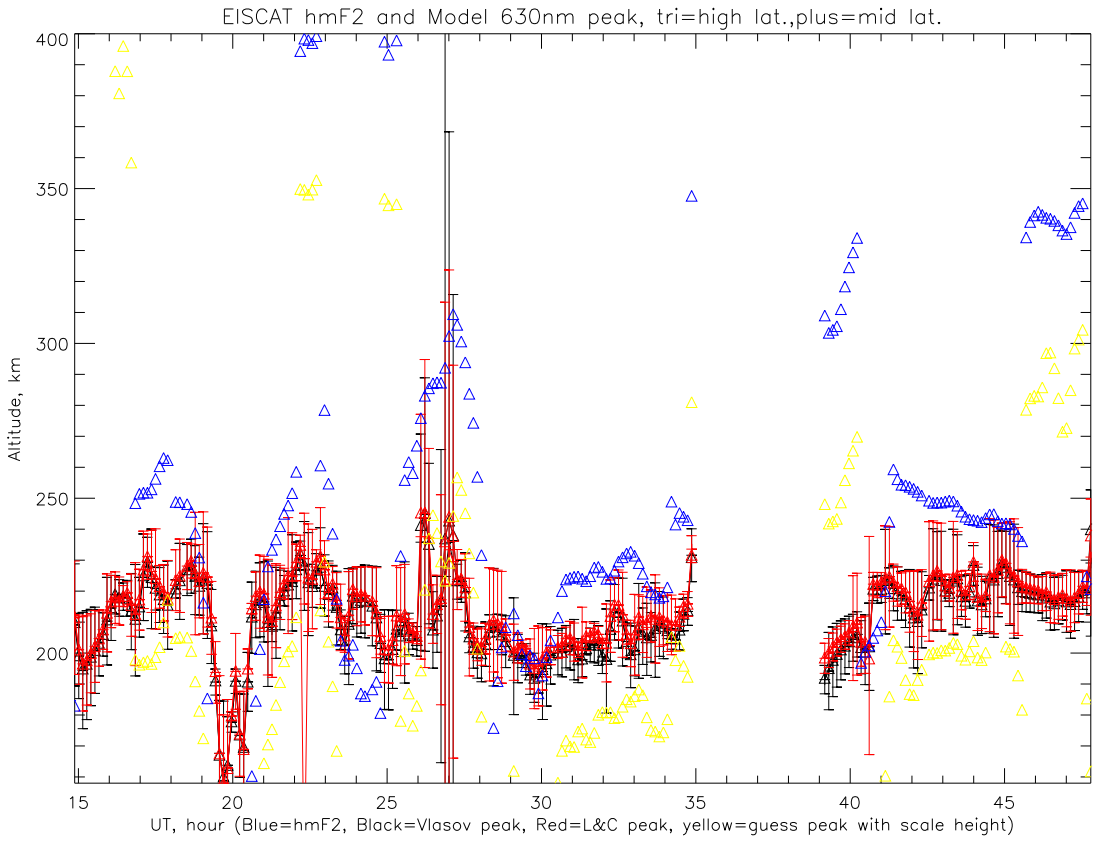


Figure 4.10: This figure shows the height of the red line emission peak determined from different methods. Red is deduced from the 630.0nm emission profiles from Link and Cogger [1988], black is from the Vlasov et al. [2005] emission profiles. Blue is the altitude of the electron density peak, hmF2. Yellow is hmF2 minus one scale height ($k_B(T_e+T_i)/m_ig$). There are also a red and black lines plotted with plus-symbols, but these are mostly overlapped by the triangle symbols. The plus-symbol represents the height deduced from photochemical models that includes $P_4^{630.0}$, and the triangle symbols without $P_4^{630.0}$.

4.2.3.8 Red line model discussion and summary

- The Link and Cogger [1988] model appears to be the slightly better model for this experiment, in terms of both magnitude and correlation match with MSP-measured intensity. It gives (model/measurement) ratio of 0.94, and correlation of 0.53.

- This suggests that the Link and Cogger [1988] model's choice of the lower Einstein transition coefficient, A_{6300}=5.15×10^{-3}, and the exclusion of quenching by atomic oxygen gives a better model-measurement match for our case.

- The Link and Cogger [1988] method also gives a slightly better correlation. The main difference between Link and Cogger [1988] and the other models is their choice of rate

coefficients, which in turn affects the smaller time scale variation in the calculated O^{2+} density, resulting in the slightly better correlation. However, the improvement is not significant.

- In terms of the altitude of the red line emission peak, there is a negligible difference between the models from the different authors, and between the high and middle latitude production processes. So the difference between the models has no significant effect in terms of FPI observation height determinations.

- The emission altitude determined from the photochemical model is supported by comparisons with EISCAT hmF2 measurements and the altitude of the electron density peak.

- The modelled emission height is also able to reflect the lower average peak red line emission altitude at 210km, appropriate for solar minimum conditions.

- The study here is a test on the height-integrated intensities and their comparisons with MSP data. So it is not possible to give further insight into the detailed shape of the emission profile. However, in terms of accompanying FPI/SCANDI experiments, the simple Link and Cogger [1988] model should be sufficient to provide a general average shape of the 630.0nm emission profile and its approximate peak altitude. Please note the author is not proposing a precise peak altitude can be determined using this method, which is also not required for the science of this thesis.

4.2.3.9 Emission model application to solar cycle trend variation

Finally, the average red line emission profiles are determined with samples of EISCAT and MSIS data listed in table 4.2. The results are shown in figure 4.11. It should be noted that the values plotted are averaged over a typical EISCAT experiment period only. Hence, they are small winter samples giving a rough representation of variation over this solar cycle.

- The top panel in the figure demonstrates that the average red line emission peak altitude (black curve) follows the altitude variation of a fixed pressure level (blue curve). The results are determined from the simple Link and Cogger [1988] model. The black curve is determined by modelling and averaging the red line emission profile for each of the cases over the solar cycle. The data entered into the red line

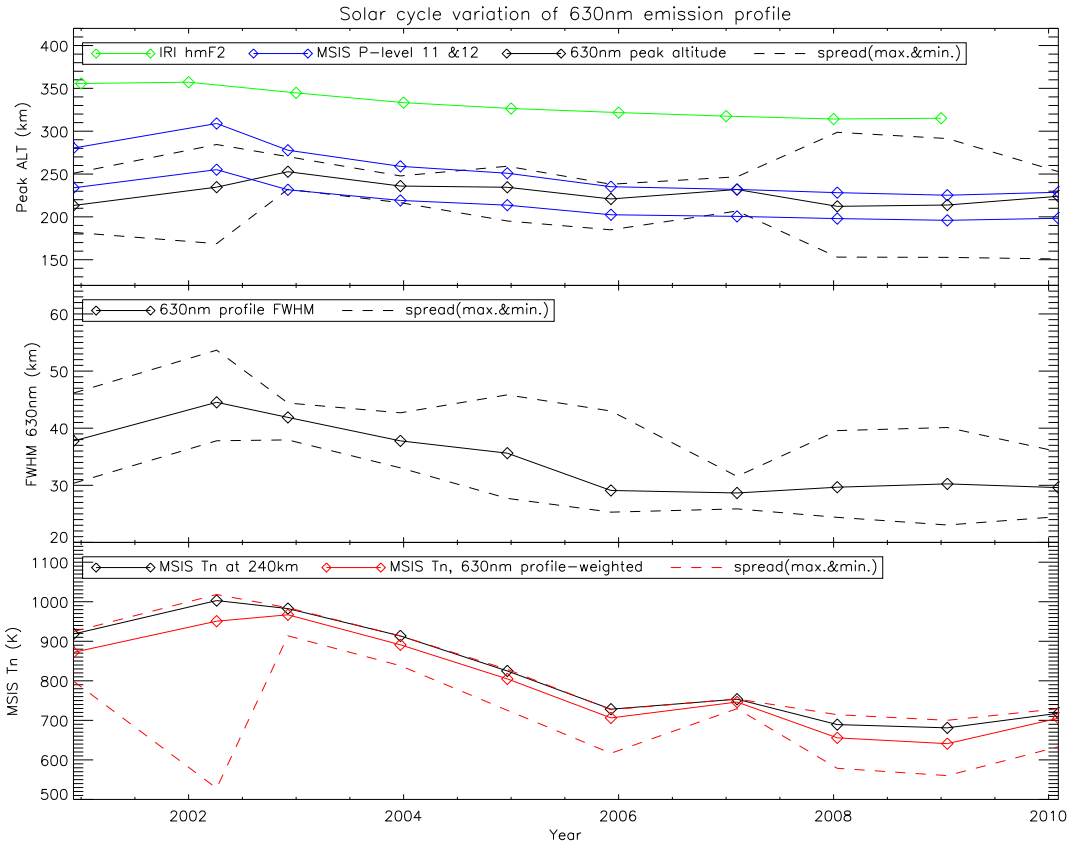


Figure 4.11: The top panel shows the IRI hmF2 altitude in green, the MSIS pressure levels 11 and 12 in blue, the red line peak emission altitude in black, and the spread (maximum and minimum) of the red line peak in dash lines. The middle panel shows the full width half maximum of the red line profile and its spread. The bottom panel shows the MSIS neutral temperature at 240 km altitude in black, the MSIS neutral temperature weighted by the red line profile in red, and its associated spread in dashed-lines. Each data point represents the average value, over an experiment in the associated year, within this solar cycle. Each experiment covers approximately one to two nights. All experiments are taken at the high latitude Svalbard site.

model are listed in table 4.2. The EISCAT electron density is the input parameter that has the highest influence on the red line model peak emission altitude (i.e. the black curve) and the red line profile shape. The fixed pressure level purple curves are simply the altitudes of two pressure levels closest to the F-region, determined from MSIS. The hmF2 variation (green curve) is also a good indicator of the average height change. The maximum peak altitude is about 240km at 2003, two years after the solar minimum because it is a high latitude site. The two-year separation of the peaks is because Earth's axis is tilted, and maximum activity on the Svalbard high

latitude site is therefore delayed. The delay is well demonstrated by plotting F10.7 (representing solar activity) and Ap indices (representing geomagnetic activity), as shown in figure 2.7. The figure demonstrates the same two-year delay between F10.7 and Ap peaks. Overall, the average red line peak variation follows the thermal expansion and contraction of the atmosphere closely.

- The middle panel shows the average full width half maximum (FWHM) of the modelled red line profile and its variation over the solar cycle. The FWHM are determined from the same data set and using the same volume emission rate profile model as the first plot. It shows that the emission profile is the widest between 2002/2003 at solar maximum, and reaches a smaller width starting from 2006. This is because on average at solar maximum, the electron density profile is wider on the lower altitude side. The electrons have a greater spread of energies and penetrating altitudes. The opposite is true for solar minimum.

- Finally, the bottom panel shows the MSIS temperatures at a fixed altitude of 240 km from the MSIS temperature profiles in black. The same MSIS temperature profiles, but weighted by the modelled red line profiles are shown in red. The red line profiles used for the temperature weighting are also determined using the same data set as the other two plots. The weighted temperature is therefore a rough imitation of the height-integrated FPI measurements. Using the model-weighted temperature, it can be seen that at solar maximum, the fixed height neutral temperature overestimates the neutral temperature by about 50K. This is because the wider red line profile has included more of the lower altitude temperature in the weighted results. The fixed altitude neutral temperature also overestimates the neutral temperature at solar minimum, but the reason is because of a drop in averaged emission peak altitude.

The results here clearly demonstrate that for any long term temperature trend study using the FPI neutral temperature, it is important that the variation in emission height, as well as the width and shape of the emission profile, are considered. The study here presents the first application of red-line volume emission rate profile, as a weighting function for FPI-measured temperatures. For any long-term neutral temperature trend studies which looks for small temperature trends at the upper atmosphere, the 50K uncertainty introduced by changing emission profile needs to be accounted for accurately.

4.3 Statistical FPI neutral temperature study, years 1999-2008

4.3.1 Introduction

Ten years of FPI temperatures from years 1999 to 2008 are investigated. These are upper thermosphere (∼240km) temperatures deduced from the 630nm emission at Svalbard from two different instruments. The instruments are the FPI and SCANDI, and they only have a very small difference in observing position. SCANDI is an all-sky imager and the FPI has a 1 degree field of field that falls within the SCANDI's all-sky detection region. The study is divided into two major sections, where one section looks at the daily average temperatures, and the other analyses the diurnal hourly-average temperatures.

In previous long-term temperature trend studies, ion temperature measurements from incoherent scatter radars, e.g. Holt and Zhang [2008], have been the preferred method. This is because radar measurements are not affected by clouds or sunlight, and they tend to have been in operation for more than three solar cycles. The ion temperature data sets are therefore larger and more continuous. The FPI is an optical instrument and can only measure in the winter for the Svalbard latitude, and clear-sky conditions are usually required. There is also no ambiguity in the altitude of the ion temperature measurements, whilst the FPI measurements are at constant pressure level and vary by approximately one scale height between solar maximum and minimum. The emission profile effect on the FPI temperature is discussed in detail in the emission profile section (section 4.2.3.9).

For the ion temperatures, T_i, experiments, T_i are assumed to be a good approximation to the thermospheric neutral temperatures, T_n, especially for geomagnetically quiet conditions. The relationship between ion and neutral temperatures are mostly governed by the simplified energy balance equation. The full equation is discussed in detail by Schunk [1975], and has been simplified by e.g. McCrea et al. [1993], Lockwood et al. [1993]:

$$T_i - T_{eq} = \frac{m_n}{3k} (V_i - V_n)^2 , \tag{4.38}$$

where T_i is the ion temperature, m_n the neutral mass, k the Boltzmann constant. T_{eq} is the ion temperature without frictional effects, and

$$T_{eq} = T_n + \Delta T_{ie}, \tag{4.39}$$

where T_n is the neutral temperature, and ΔT_{ie} is the electron heating by thermal

contact with the neutral atmosphere. The relationship between T_i and T_n is a good indicator of the level of ion-neutral coupling. It should be noted that equation 4.38 applies to the 'true' thermodynamic ion temperature, but this does not necessarily match the measured ion temperature under extreme conditions when, for certain species such as O^+, the ion temperature can become anisotropic.

Although ion temperature measurements for long term atmospheric temperature studies is the usual preferred method; there is the possibility that the neutral temperatures are not represented well by ion temperatures. The discrepancy may even be present at quiet times, where V_n and V_i usually differ by 100m/s or so [Lockwood et al., 1993]. A time delay between an ionospheric event and the neutral response is also expected, determined by the inverse of the ion-neutral collision frequency. It is also named the e-folding time (section 2.3.4). The time delay could have significant influence on long-term neutral temperature trend studies, which look for trends as small as -4.7K per year [Holt and Zhang, 2008]. Hence, direct neutral temperature measurements can provide further insight into long-term temperature trends, and may give a different trend result if there are any statistical differences between the ion and neutral temperatures.

The ten years of FPI data chosen for the study here are of insufficient length to give any long term temperature trend results, but previous ground-based FPI temperature studies are even shorter or event-based [Smith and Hernandez, 1995, Killeen et al., 1995]. Also this is the first long term FPI neutral temperature trend study available for Svalbard to the author's knowledge, making it a scientifically valuable high latitude data set. The changes in the red-line airglow model measured by the FPI emission profile are also considered carefully and their effects on the long-term FPI-measured temperature are estimated. The application of the airglow model to a long-term trend study is also new.

4.3.2 Solar cycle trend

Even though the FPI data analysis described in section 4.2 concludes that the daily average data are not significantly affected by cloudiness, the data used for the temperature trend study here are restricted to clear-sky data only. This is for consistency because only clear sky data is suitable for the diurnal temperature variation analysis. Hence the same data set is used for the daily averaged data here and the UT diurnal-averaged data in the following section. However, for any future long term temperature study that applies daily-average results, the clear-sky data do not need to be isolated.

The FPI temperature data are carefully calibrated with an instrumentation function, and/or compared with SCANDI data when SCANDI is in operation. The only exception is 2005 when neither are available, and hence the calibration function from the previous year is used.

4.3.2.1 Determination of the most appropriate sampling method

The Svalbard FPI observing season usually begins in October, and ends in March in the following year. However, due to sunlight, it is only in operation for full or nearly a full 24 hours from November to January. The typical distribution of data points for different UT over a full observing season can be found in figure 3.20. All sampling methods use a simple mean to calculate the average values. In order to have the largest possible data set, two sampling methods are tested:

- Data set A: Using all 24 hours of data, but restrict the data to November, December and January data only. Hence the sample only contains data from mid-winter, but any diurnal variation will be averaged out, which is the only suitable sampling method if clear-sky data are not isolated.
- Data set B: ± 3 hours from local midnight are chosen (21 to 03 UT), for all six months of each FPI observing season. This method restricts the FPI data to night time data only, but covers data from October to March.

The data are further divided according to activity with local Longyearbyen K indices, which are deduced from the University of Tromsø magnetometer readings. The data are also separated according to the global Kp indices. Each of the FPI-measured data points are then binned according to geomagnetic activity. Both the global Kp indices and local K-indices are at 3 hours time resolution, and they are separated into the following activity groups:

- ALL - includes all data points, irrespective of local K or global Kp indices.
- QUIET - Global Kp indices = 0-2. Local K indices = 0-2.
- MODERATE - Global Kp indices = 3-5. Local K indices = 3-5.
- ACTIVE - Global Kp indices = greater than 5. Local K indices =greater than 5.

Fully calibrated SCANDI data from season 2007/08 are chosen to test the effect of the different sampling methods, as well as the differences between sorting activity level with global or local K indices, or the solar activity $F_{10.7}$ values. The differences between the two different sampling methods are shown in figures 4.12 and 4.13. They show the '12-11UT January and December only' data set (A), and the '21-03UT October to March' data set (B) respectively. The activity levels are separated by local K indices for both figures.

The main difference between the two sampling methods is that data set A gives an average temperature that is lower by about 50K than data set B, if the moderate and active data groups are ignored because they are very small samples. It indicates that the mid-winter average which covers all 24 hours, is lower than the night-time only observing-season average. This result is expected because the hours chosen as night-time (21-29UT) in method B cover one of the peaks in the diurnal neutral temperature variation. The diurnal peaks are shown in e.g. figure 4.5. One of the peaks is found in the mid-night sector, and the other around the cusp.

The mid-night sector peak is modelled by CMAT2 in figure 4.24. The common cause of the midnight peak not being noted is because the overall wind is too large in most thermospheric models. The reason is most existing models do not allow for variation in the electric field and assumes steady state wind. Hence, the wind is too high and the cooling term becomes too large. Furthermore, 2007/2008 is part of the extremely low activity solar minimum and the auroral oval was possibly unusually contracted, contributing to the higher mid-night peak.

Furthermore, the two months of data from data set A gives a larger sample size for this particular year, since it covers the full 24 hours. Hence, data set A is the preferred method for a temperature trend study, and does not create a bias towards any features in the diurnal temperature variation. It also gives a larger sample with the full 24 hour coverage, and any unaccounted for cloud effects are more likely to be averaged out (as shown in section 4.2.2).

4.3.2.2 The solar cycle temperature trend

Only data from November, December and January of each year (data set A) are used. The data-set is restricted only by instrumentation maintenance downtime, availability of calibration function and cloudiness (unless day-averaged is applied). The daily-averaged neutral temperatures are shown in figure 4.14.

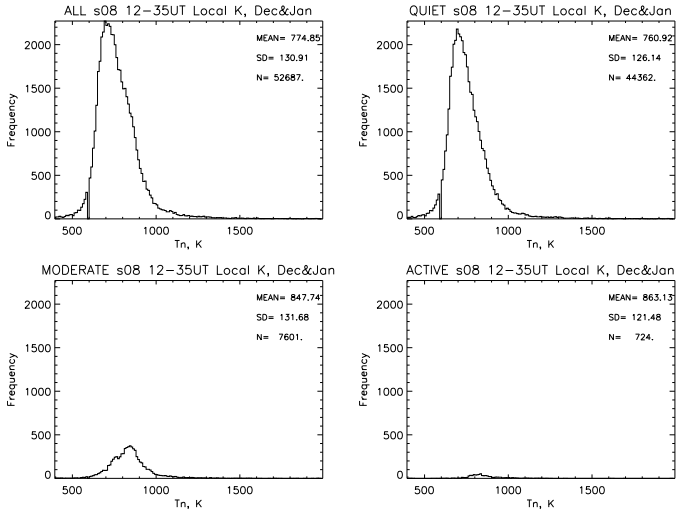


Figure 4.12: Data set (A). This figure shows the SCANDI-measured neutral temperature distribution, for temperatures in December and January only for season 2007/08. All 24 hours of data, as well as all 25 SCANDI observing directions are included. The data are divided according to geomagnetic activity with local K indices. The top left plot includes all data points irrespective of activity. The top right plot shows the quiet data (K=0-2), bottom left is the moderate data (K=3-5) and the bottom right is the active data (K>5). The mean, standard deviation (SD) and the number of data points (N) are also shown in each plot.

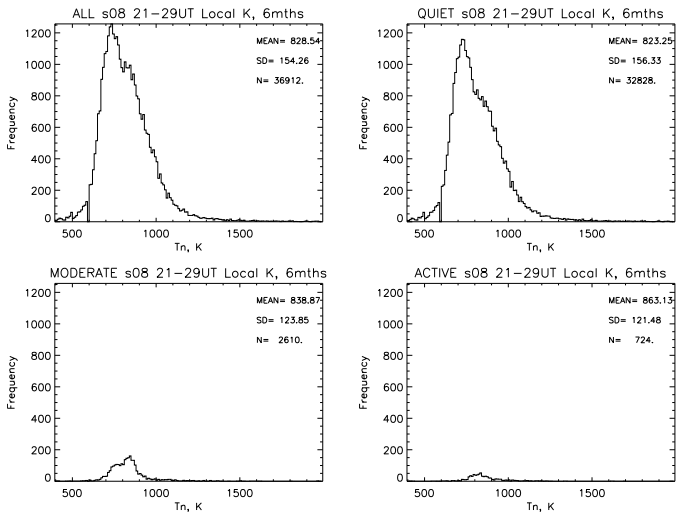


Figure 4.13: Date set (B). This is very similar to the figure above (figure 4.12), the only difference is that temperature data from all six months (Oct-Mar) of the observing season are included, but are restricted to night-time data (21-03UT) only.

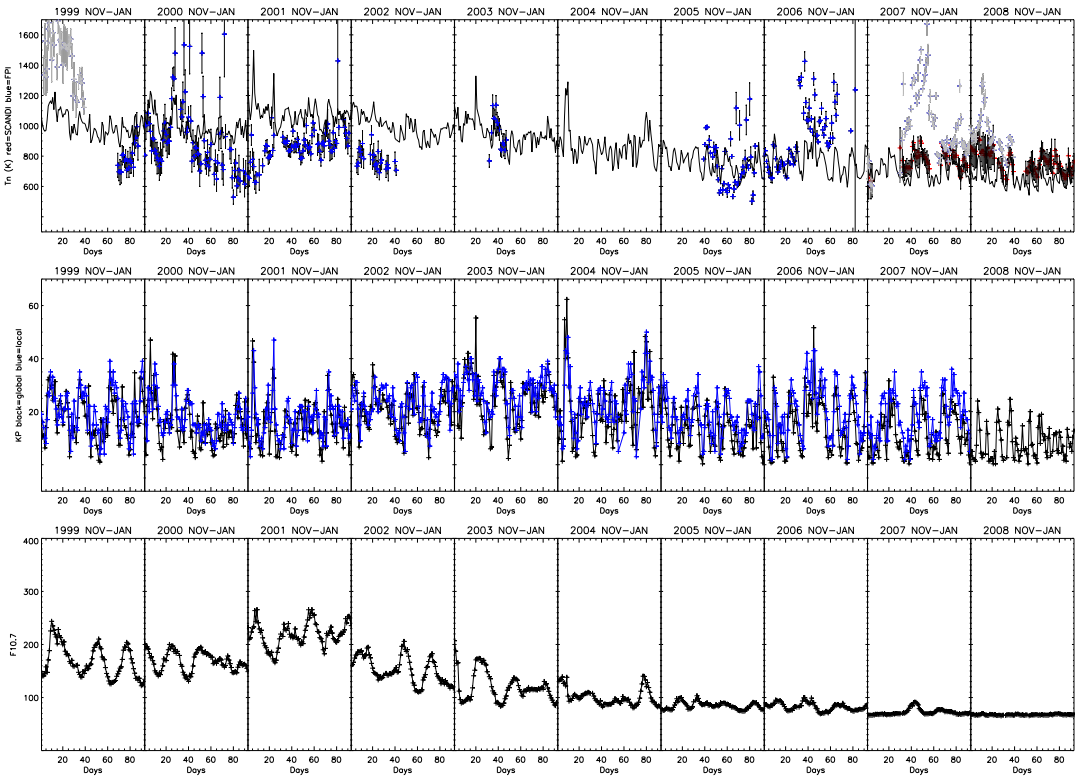


Figure 4.14: Daily-averaged neutral temperatures over the solar-cycle are plotted in the top panel. Each box in the panel represents November to January for each year. The temperature data are also averaged over all the FPI or SCANDI directions. The black line is MSIS data, the blue data points are FPI-measurements, and the red data points are SCANDI-measurements. Unsuitable data points because of calibration problems or moon-lit conditions are greyed-out. The second panel is the geomagnetic activity level, with the daily-sum local K index in blue, and the global Kp indices in black. The bottom panel is the daily $F_{10.7}$ values.

It is clear that data gaps are present. However, even for the smallest data set in 2003, there are 7218 data points in the sample, including all FPI observing directions. The data points that are excluded (greyed-out in the figure) are a) points that are affected by moon light, which have a distinctive variation that follows the moon phase, or b) those without a suitable temperature calibration function. Normally, one calibration profile is taken for the FPI per observing season, and the dates at which the profiles are taken are shown in table 4.3. For example, the elevated November temperatures in 1999 are likely to be unreliable because the calibration profile was taken in December 1999. It is expected that the data are more reliable if they are obtained close to a calibration, which corrects for error such as instrumentation temperature drift. Note that from 2007 onwards, SCANDI data are chosen because it self-calibrates with approximately 4 calibration profiles taken

per night.

Table 4.3: Laser calibration profile for each year.

Season (Year)	LASER profile date
1999	10^{th} December
2000	28^{th} October
2001	16^{th} November
2002	15^{th} January
2003	2^{nd} December
2004	NO DATA
2005	NO CALIBRATION
2006	5^{th} December
2007	OK - SCANDI data (self calibrates)
2008	OK - SCANDI data (self calibrates)

Figure 4.14 also shows the global Kp indices and the local K index. They have similar variations, but the local K index has higher fluctuations because of local geomagnetic activity. The $F_{10.7}$ values in the figure indicate peak values in 2001, which is different from the time when peak values of K and Kp indices are found in 2003.

An average is then taken of the temperature data available in the three months period for each year to represent the average winter temperatures over the solar cycle, and are shown in figure 4.15. The FPI-measured temperature (black line, top panel) has an extended peak from 2001 to 2004. However, the modelled MSIS temperature shows a distinct peak in 2002 at fixed 240km altitude (blue line, top panel).

The ion temperatures are shown in the middle panel in black. The data here are one-night average samples only, and are measurements from the Svalbard EISCAT ESR radar (see table 4.2). At solar maximum years, the one-night T_i has an average value of $\sim$1130K, while the MSIS T_n gives an average of about 950K, and the FPI gives an average of around 850K in the top panel. T_i is higher than both measured and modelled T_n. The approximately 200K higher T_i is expected and acceptable. This is because ion temperature in general is expected to have a temperature higher than that of the neutral temperature, before it is completely thermalised through ion-neutral collision, depending on the e-folding time [Heelis et al., 2002, Kosch et al., 2010]. However, it is unexpected that the MSIS T_n is lower than the measured T_n.

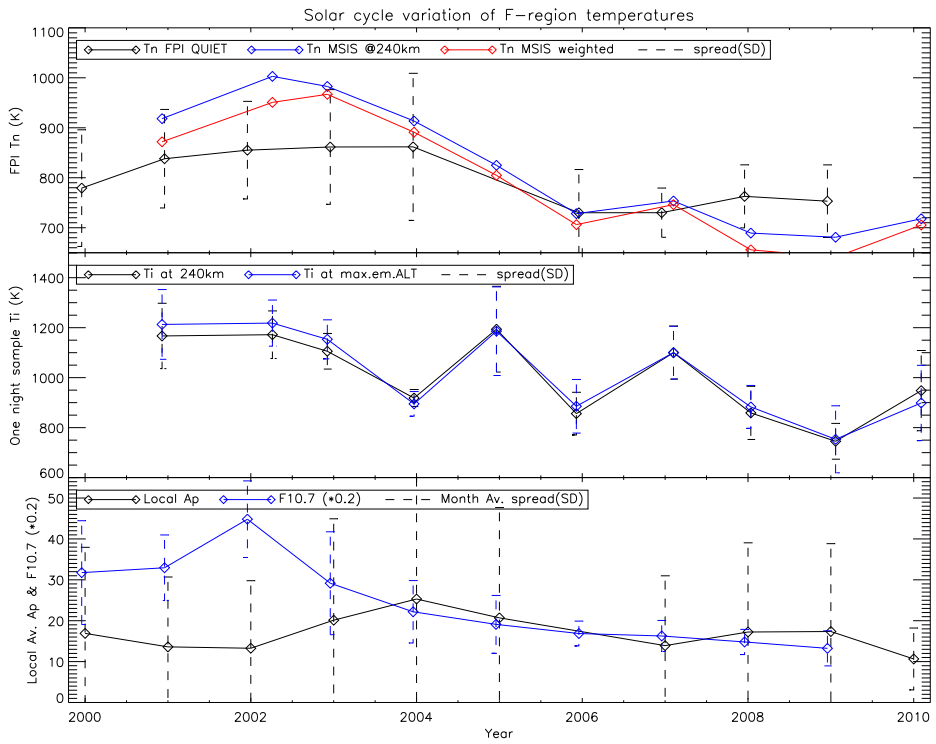


Figure 4.15: The top panel shows the winter-averaged (November, December, January) temperature data for each year, over the solar cycle. The black graph is the FPI/SCANDI measured neutral temperatures, the blue graph is the MSIS modelled temperature at fixed 240km altitude, and the red graph is also the MSIS modelled temperature, but is weighted by the red line emission profile. The middle panel shows the EISCAT ion temperature at fixed 240km altitude in black, and the ion temperature at the peak red line emission altitude in blue. The bottom panel shows the average winter AP values in black, and the average $F_{10.7}$ value multiply by 0.2 in blue.

At solar minimum, T_i has a good match with measured-FPI T_n at ∼750K, but the MSIS-modelled T_n gives a lower value of about 600K. It is usually assumed that T_i is a good approximation to T_n during quiet conditions, which is well demonstrated by the T_i-FPI T_n comparison here.

Comparing the measured and modelled T_n with the averaged $F_{10.7}$ and local Ap values, $F_{10.7}$ has a distinct peak in 2002. However, the local Ap peaks in 2004. This result matches with the expectation that the maximum direct sun-spot contact at high latitude appears a few years after solar maximum, due to the time it takes for the area where sunspots are created to migrate to the solar latitudes that have the most influence on the Earth. (figure 2.9).

4.3.3 Solar maximum and minimum comparison

The solar maximum neutral temperatures in 2002, and solar minimum results in 2008 are looked into further, and the measured and modelled T_n distributions are shown in figures 4.16 and 4.17. The MSIS data are taken directly from the official MSIS website, which automatically generates a value for every 15 minute. The MSIS model was not run locally in UCL by the author. Therefore the sample size is different.

For the solar minimum data (figure 4.16), the difference (T_{FPI}-T_{MSIS}) between averages from all data points is 58±2K; from quiet geomagnetic conditions only is 59±3K; from moderate conditions only is 113±3K, and from active conditions only is 106±8K. These are similar to previous studies that compare FPI measurements and MSIS modelled temperatures e.g. Killeen and Roble [1984], Smith et al. [1998], where MSIS underestimates neutral temperatures under high activity conditions. Furthermore, the spread of the measured data is at least double that of modelled data for all geomagnetic conditions. A wider measured peak indicates a wider spread of energy input to the thermosphere.

For the solar maximum data (figure 4.17), it is surprising that the measured temperatures are lower than the modelled temperatures. There is always the possibility that this is because of poorly calibrated measurements, but not in this case because a calibration profile was taken on 16th November, 2001 (table 4.3), for the three months of data (November 2001, December 2001, January 2002) in this year. It is unlikely that the calibration measurement using a laser (chapter 3) would be incorrect. Also the temperature distribution in figure 4.17 is smooth, and without any unusual jump or drop in the T_n measurements that might indicate calibration error. Therefore, there is no reason to suggest the measurements are not accurate.

For the solar maximum results, the temperature differences (T_{FPI}-T_{MSIS}) are -229±5K for the all data points average, -235±3K for quiet activity average, -238.21±4K for moderate activity average. There are no MSIS data available for the active period from the MSIS data base at the time of this thesis, but the measurements give a higher temperature of 152±8K than the quiet only average. Overall, the measurements here show that MSIS T_n is systematically higher than measurements by ∼230K. The smaller sample of ion temperatures for this year give a value of ∼1200K, which is higher than both the modelled and measured T_n, so it does not provide a useful comparison.

The sample size of the ion temperature is listed in table 4.2. They are only sufficient as rough comparisons with FPI temperatures, but not as a statistical study of ion

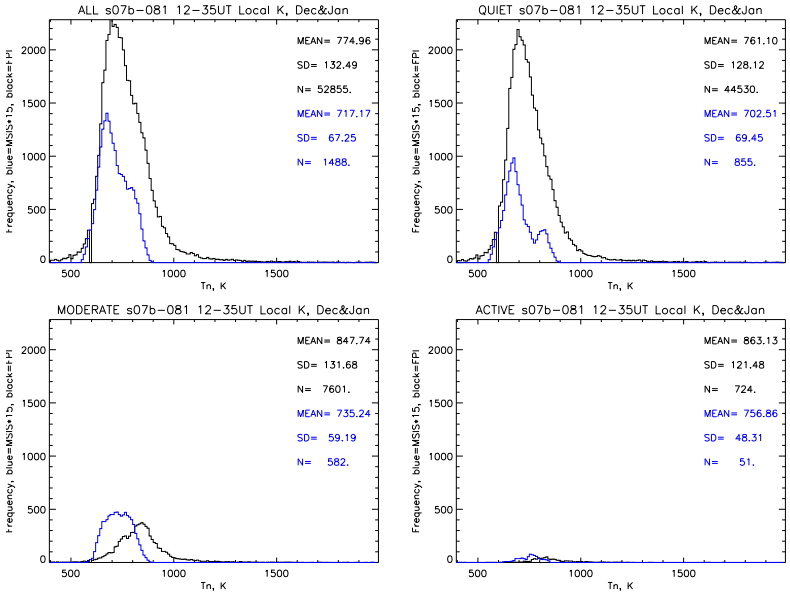


Figure 4.16: This figure shows the SCANDI-measured neutral temperature distribution, for temperatures in December and January only for season 2007/08 (solar minimum) in black. These are F-region red line SCANDI data. The MSIS modelled temperatures are shown in blue. This result is similar to figure 4.12. All 24 hours of data, as well as all 25 SCANDI observing directions are included. The data are divided according to geomagnetic activity with local K indices. The top left plot includes all data points irrespective of activity. The top right plot shows the quiet data (K=0-2), bottom left is the moderate data (K=3-5) and the bottom right is the active data (K>5). The mean, standard deviation (SD) and the number of data points (N) are also shown in each plot.

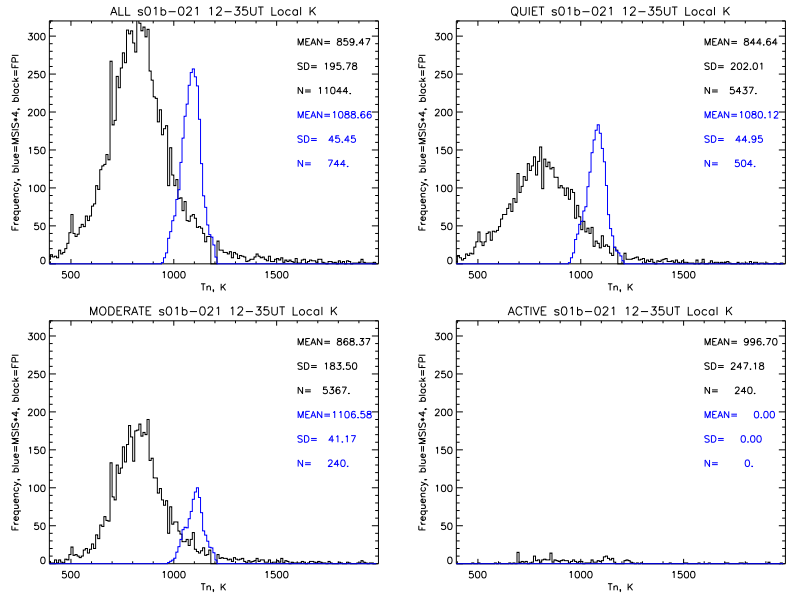


Figure 4.17: This is very similar to the figure above (figure 4.16), the only difference is that temperature data from season 2001/02 (solar maximum) are shown.

temperatures on their own. The focus of this thesis is the FPI neutral temperature.

The systematically higher and lower daily-averaged measured T_n than modelled T_n, for solar minimum and maximum respectively, are shown clearly in the T_n measured vs. T_n modelled scatter plots in figure 4.18. The measured and modelled T_n in general are well correlated with each other, especially for the solar minimum half of the solar cycle. The exception is 2000, when measurements show a much larger spread of value than MSIS. The 2001 solar maximum plot clearly shows systematically lower values than modelled T_n for all temperatures, while the opposite is true for solar minimum e.g. year 2006. Note that the systematic difference is larger for solar maximum.

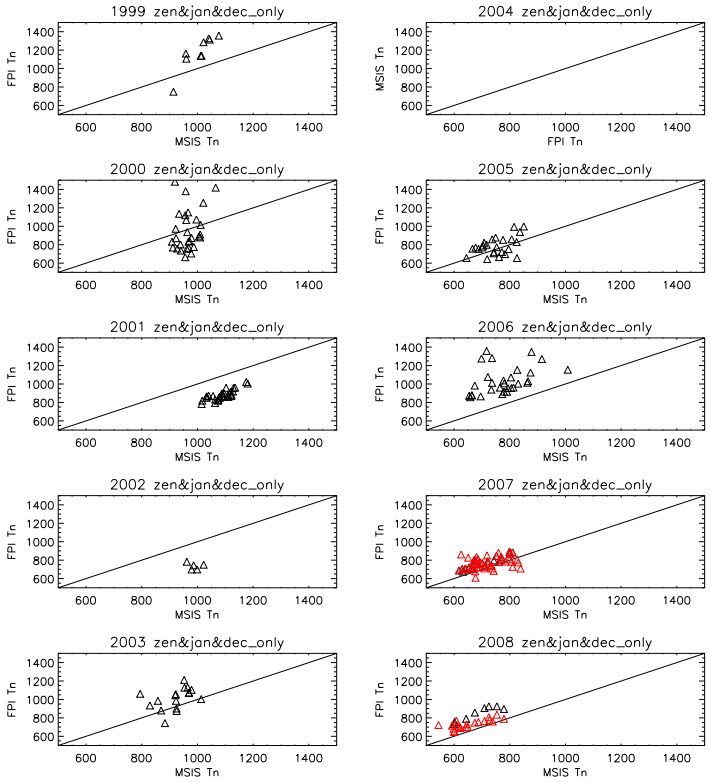


Figure 4.18: Each plot shows the daily-averaged FPI-measured T_n vs. MSIS-modelled T_n for December and January, from the zenith observing direction only. The plots in the left column are for years around solar maximum 1999 to 2003 from top to bottom, and the right column are for years around solar minimum 2004 to 2008. There are no 2004 data. Red symbols indicate SCANDI data.

MSIS modelled T_n that are lower than measurements have been found before, usually because MSIS's use of $F_{10.7}$ as the solar flux proxy may not always be representative of the Sun's ultraviolet irradiance [Aksnes et al., 2007]. Furthermore, MSIS is an empirical model based on years of observations, but temperature measurements of the middle thermosphere

are sparse between 130 and 250km altitude. For example, Aksnes et al. [2007] shows that the percentage match between model and measurement can be improved if the MSIS temperature is lowered by 43K for a June 2004 satellite data set taken at 163km altitude. So although higher than modelled T_n measurements are expected, there are sound reasons for why MSIS may over estimate T_n.

Other possible explanations for the higher modelled T_n include the high correlation of MSIS T_n with solar activity ($F_{10.7}$), where $F_{10.7}$ peaks in the same year as the MSIS T_n in 2002; but the solar maximum FPI T_n shows a much less intense extended peak (figure 4.15), possibly more correlated with Kp. However, the Kp indices do not show a matching extended peak, so the MSIS-FPI T_n difference may not be simply explained with a better correlation between T_n and Kp indices. The correlation of T_n with $F_{10.7}$ and Kp are explored in section 4.3.4.

4.3.3.1 Emission profile effect on the Tn trend results

The effect of solar variations in the red line emission profile on the FPI-measured temperatures are presented in the data preparation section (section 4.2.3). It shows the general effect is that the FPI-measured temperature is lowered by approximately 50K at both solar maximum and minimum. The causes are different: at solar maximum, the lower temperature is because of the wider red line profile width on the lower altitude side. The FPI measurement therefore includes more of the lower altitude contribution and results in lower temperatures. At solar minimum, the lower temperature is caused by the lower peak red line emission altitude.

Therefore for the solar minimum results, the changing emission profile is unlikely to be the reason for T_{FPI}-T_{MSIS}=58±2K, where the FPI temperature is expected to be even higher if the emission profile effect is taken into account. The emission profile effect also further supports that for the Svalbard site, where the average neutral temperature is not unusually low for the extended solar minimum period. However, for the solar maximum results, where T_{FPI}-T_{MSIS}=-229±5K; it is possible the emission profile change can partially explain the FPI-MSIS discrepancy. Hence the FPI temperature could in fact be ∼50K higher because of the wider red line profile width on the lower altitude side. The T_{FPI}-T_{MSIS} difference is then reduced to around -180K. Nevertheless, the difference between measurement and model at solar maximum remains high.

For both the solar minimum and maximum cases, the FPI data have been carefully

analysed, with consideration to instrumentation calibration, its diurnal variation and relationship to thermospheric model, different sampling methods, effect of clouds on the data, and finally the effect of the changing emission profile. The effect of the emission profile on FPI measured data is also a new method that has not been considered previously to the author's knowledge. Possible causes of the discrepancy between the FPI and MSIS temperatures are also investigated further by comparisons with solar and geomagnetic proxies, such as F10.7, S10.7, Kp and AE indices, as shown in the sections that follow.

4.3.4 Neutral temperature correlation with solar and geomagnetic proxies

The temperature, density and composition of the Earth's upper atmosphere are the most sensitive to the EUV spectral range of the solar irradiance. There are a number of indices available to represent the solar EUV variation, and a good description of the solar EUV is required to define accurately solar EUV forcing in upper atmospheric models. The most common solar proxies are $F_{10.7}$, $S_{10.7}$ and MgII (section 2.5.1). Studies such as Dudok de Wit and Bruinsma [2011] and Floyd et al. [2002] investigate the most appropriate proxy for thermospheric density modelling, and the conclusion is that $S_{10.7}$ produces the most correlated results. The FPI temperatures here are compared to both $F_{10.7}$, which is the solar proxy chosen by MSIS, and the newly proposed solar proxy $S_{10.7}$. The difference in correlation between the two proxies are explored. In addition to $F_{10.7}$ and $S_{10.7}$, the FPI temperatures are compared to two geomagnetic indices appropriate for the high latitude site, and they are the Kp/Ap and AE indices (section 2.5.1).

Also as suggested by Killeen et al. [1995], the correlation with $F_{10.7}$ demonstrates a link between neutral temperature and solar forcing, while the correlation with Kp demonstrates the relationship between neutral temperature and geomagnetic activity. Therefore the correlation results can give further information on the balance of solar forcing and the geomagnetic activity effect on thermospheric temperature over this solar cycle.

4.3.4.1 T_n vs $F_{10.7}$ and T_n vs Kp, daily average

The correlation results are investigated with data from the previous solar cycle provided by a different study: Killeen et al. [1995] use 6 years of FPI data in the polar cap region at geomagnetic latitude 76.5N at Thule, Greenland. The Longyearbyen site is located at 75.0N geomagnetic latitude. Therefore the two sites are comparable geomagnetically.

Killeen et al. [1995] describe a higher T_n dependence on solar activity than geomagnetic activity in their statistical results. Their conclusion is based on a better correlation of neutral temperature with $F_{10.7}$ than Kp.

The temperature data from Killeen et al. [1995] are binned according to the solar EUV flux index $F_{10.7}$, while the geomagnetic index Kp are restricted to Kp=3 only. The Tn vs. $F_{10.7}$ relationship is described by T_n=836+3.1*$F_{10.7}$ with correlation R=0.68. The data are also binned according to Kp, and with $F_{10.7}$ restricted to values between 150 and 180, and the Tn vs. Kp relationship is described by T_n=1215+7.77*Kp with correlation R=0.36. These are the results published by Killeen et al. [1995].

A similar test is then applied to the UCL FPI data in Svalbard. The data set here is also longer with 10 years of data. The FPI results are further divided into samples that include solar maximum data only (1999-2004) and solar minimum data only (2005-2009), as well as the original sample that contains data from all ten years. The Spearman's rank correlation IDL function is used, which readily outputs the rank correlation coefficient and the significance of its deviation from zero. The correlation results are summarised in table 4.4.

Table 4.4: This table shows the correlation coefficient (R) between T_n and $F_{10.7}$, and between T_n and Kp indices. The table shows the solar cycle correlation results from Thule (1986-1991) [Killeen et al., 1995], as well as correlation results from Svalbard. The Svalbard investigation is further separated into data sets including the whole solar cycle (1999-2009), solar maximum only (1999-2004) and solar minimum only (2005-2009).

Data set	R (T_n vs. $F_{10.7}$)	R (T_n vs. Kp)
Thule [Killeen et al., 1995], half solar cycle	0.68	0.36
Svalbard, whole solar cycle	$0.23 \pm 2\times10^{-6}$	$0.29 \pm 5\times10^{-7}$
Svalbard, solar max. only	$0.34 \pm 2\times10^{-4}$	0.13 ± 0.2
Svalbard, solar min. only	$0.05 \pm 4\times10^{-1}$	$0.3 \pm 2.0\times10^{-7}$

The correlation coefficients between Tn and $F_{10.7}$ for sampling from the whole solar cycle, solar maximum only and solar minimum only are 0.23, 0.34 and 0.05 respectively;

and are described by T_n=(707±2) + (0.82±0.02)*$F_{10.7}$, T_n=(581±16) + (1.3±0.08)*$F_{10.7}$, and T_n=(755±5) + (0.19±0.06)*$F_{10.7}$. The Kp values are restricted to the range 0-20 for the T_n-$F_{10.7}$ correlation test. The correlation coefficients between T_n and Kp for sampling from the whole solar cycle, solar maximum only and solar minimum only are 0.29, 0.13 and 0.27 respectively; and are described by T_n=(710±0.8) + (7.1±0.05)*Kp, T_n=(728±10) + (3.2±0.48)*Kp, and T_n=(721±0.8) + (5.91±0.05)*Kp. The $F_{10.7}$ values are restricted to the range 65-125 for the T_n-Kp correlation test.

The general conclusion is that the overall correlations between T_n and both indices are low in this solar cycle. The highest correlation coefficient of 0.34 is found by the T_n vs $F_{10.7}$ correlation. The correlation is significantly lower than the 0.68 value found by Killeen et al. [1995] in the previous solar cycle. There is another unique result from this solar cycle and the extended solar minimum: the variation of $F_{10.7}$ is extremely small in the solar minimum sample, ranging only between approximately 65 and 90. This is the cause of the near 0 correlation for T_n vs. $F_{10.7}$ in the solar minimum only sample, as well as the small improvement in correlation from 0.23 to 0.34 when only solar maximum data are used. Overall, the correlation remains too low to be significant. The results from the usually weak solar cycle and extended solar minimum give a different correlation conclusion from that by Killeen et al. [1995].

4.3.4.2 T_n vs $S_{10.7}$ and T_n vs AE, daily average

Following studies by Dudok de Wit and Bruinsma [2011] and Floyd et al. [2002], which investigate the most appropriate proxy for thermospheric density modelling; the daily-average FPI neutral temperature correlation tests are performed for T_n vs $S_{10.7}$ and T_n vs AE (Auroral Electrojet index). Information about $S_{10.7}$ and AE are available in section 2.5.1. $S_{10.7}$ is chosen because it represents the wavelength band of 26-34nm in the Solar EUV Monitor (SEM), and Dudok de Wit and Bruinsma [2011] find that it systematically gives a more accurately modelled thermospheric density than $F_{10.7}$ and MgII, when the modelled results are compared to 14 years of daily-mean density measurements. Their results are deduced from global positions determined by the French geodetic Stella satellite in a 96^0 inclination and near-circular orbit. Their data are also taken at a much higher altitude at 813km than the FPI red-line data. Solar forcing proxy is the focus of their experiment, and the authors note that their geomagnetic proxy, Kp, also has the potential to be optimised. In the FPI study here, AE is the chosen alternative geomagnetic proxy to

be tested. It is chosen because the auroral electrojet is expected to have a direct influence on the high latitude neutral temperature, mostly through ion-neutral coupling.

The T_n vs $S_{10.7}$ and T_n vs AE correlation results are compared to T_n vs $F_{10.7}$ and T_n vs Kp results in table 4.5. Note that daily values are used in the correlation test, and the same Spearman's rank correlation IDL function is applied as section 4.3.4.1.

Table 4.5: This table shows the correlation coefficient (R) between T_n and $F_{10.7}$, T_n and $S_{10.7}$, between T_n and Kp, and T_n and AE indices. The table shows the solar cycle correlation results from Svalbard. The experiment is further separated into datasets including solar maximum data only (1999-2004) and solar minimum only (2005-2009).

Indices / proxy	R, solar max.	R, solar min.
T_n vs. $F_{10.7}$	$0.34 \pm 2\times10^{-4}$	0.05 ± 0.4
T_n vs. $S_{10.7}$	$0.29 \pm 2 \times10^{-3}$	0.09 ± 0.09
T_n vs. Kp	0.13 ± 0.2	$0.27 \pm 2.0\times10^{-7}$
T_n vs. AE	0.15 ± 0.07	$0.31 \pm 3.0\times10^{-7}$

The conclusion is that there is very little improvement between which proxy is used for correlation tests with the FPI neutral temperatures, for both the solar maximum and minimum samples.

The result here is different from the conclusion in Dudok de Wit and Bruinsma [2011]. Possible reasons include: the experiment here is applicable to the Svalbard polar cap site only and only winter data under 24 hours of darkness are used. The altitude range is also different for the FPI and is around the F-region peak between 210-260km. Furthermore, the data tested here are neutral temperatures, instead of thermospheric density. Hence, the results here are specific to the location and time for this study only, which are different from those of Dudok de Wit and Bruinsma [2011].

Nevertheless, the tests here are able to provide a pointer to neutral thermospheric modelling with respect to using different proxies: the proxies are unlikely to be a significant factor affecting the modelling of daily average neutral temperature in the polar cap. The use of different proxies is also unlikely to explain the measured and modelled temperature differences seen between the FPI and MSIS data (section 4.3.2.2). However, only two alternative proxies are tested and there is the possibility that untested and better proxies can be found. The localised results are also particularly of interest to models such as CMAT2, which is currently developing a high resolution nested grid method in order to

model the highly varying arctic thermosphere.

4.3.5 The solar and geomagnetic conditions that may cause modelled and measured results to differ, solar maximum

The aim is to understand the interdependence of neutral temperature with both Kp and $F_{10.7}$, and to discover the geomagnetic and solar conditions when MSIS is unable to accurately model F-region thermospheric temperatures. The method used is a contour plot of Ap and $F_{10.7}$, with T_n as the third variable. This is similar to the contour plots produced by Killeen et al. [1995] for temperature results from the previous solar cycle.

In a similar manner to the Killeen et al. [1995] analysis, the Svalbard FPI data are smoothed and interpolated to remove data gaps. The T_n vs. Ap values are interpolated at Ap points of 1, 4, 7, 15, 27, 48, 80, 132, 207 and 400. A two bin smooth is then performed on the T_n data. In terms of T_n vs. $F_{10.7}$, the interpolation points are 200 equal increments between the maximum and minimum $F_{10.7}$ values. A six bin smooth is then applied to the T_n temperature data.

Each SCANDI data point is first associated with a pair of Ap and F10.7 value as listed above. In the Ap plot direction (i.e. the vertical axis of the contour plots), the smoothing of Tn is smaller than the F10.7 plot direction (i.e. the horizontal axis of the contour plot). Please note the smoothing is applied independently for Tn vs. Ap and Tn vs. F10.7. Also as for the Killeen et al. [1995] method, the smoothing is to remove data gaps, which is likely to exist for a particular F10.7 and Ap pair. Furthermore, the data density plot below is able to highlight the regions of the contour plot that depend on interpolated data only, as well as regions of lower sample numbers.

The data are again divided between solar maximum (1999-2004) and solar minimum (2005-2009) samples. Both the FPI and MSIS solar maximum temperature results are shown in figure 4.19, and the equivalent data densities are shown in figure 4.20.

The solar maximum results here can be discussed with the average temperature results and distributions shown in figures 4.15 and 4.17. The main conclusion from section 4.3.3 is that the average MSIS temperature is unexpectedly higher than the FPI measured temperature at solar maximum.

The contour plots of temperature and density of $F_{10.7}$ vs Ap (figures 4.19 and 4.20) respectively demonstrate that the unusual MSIS T_n > FPI T_n is found at low solar and geomagnetic activity levels of Ap=0-80 and $F_{10.7}$=90-140. This is also the region with the

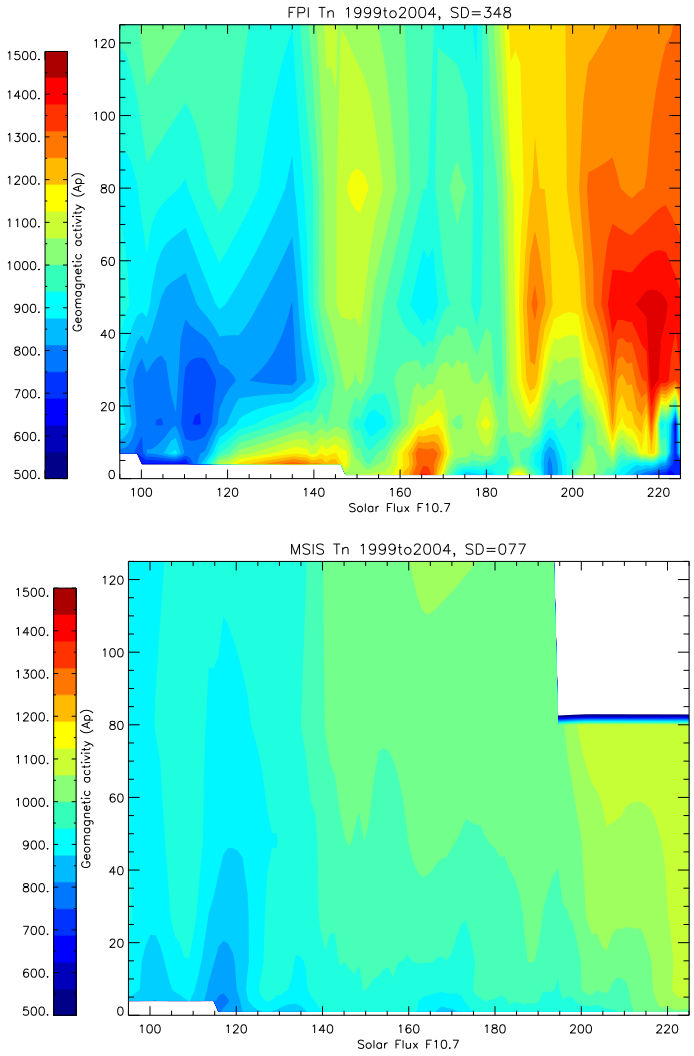


Figure 4.19: These are the hourly Svalbard FPI T_n measurements vs. Ap and $F_{10.7}$ on the top, and the equivalent MSIS T_n measurements on the bottom, for solar maximum years 1999 to 2004. The standard deviation (SD) for all the data points included in the plot are provided within the title. The colour scale unit is Kelvin.

highest data density. Therefore, when calculating an overall average, this region of data dominates the mean computation, resulting in the overall higher average MSIS T_n.

For the data density plot (figure 4.20), the density of SCANDI (and MSIS) data at a particular MSIS value are the actual number of SCANDI data points that have been used i.e. the non-smoothed/non-interpolated data. They give the true representation of the input data density. Hence, the gap in the density contour plot that is not in the temperature contour plot (figure 4.19), is a region of interpolated temperatures. Please also note the lowest data density is 88 data points, which is a significant sample size. The

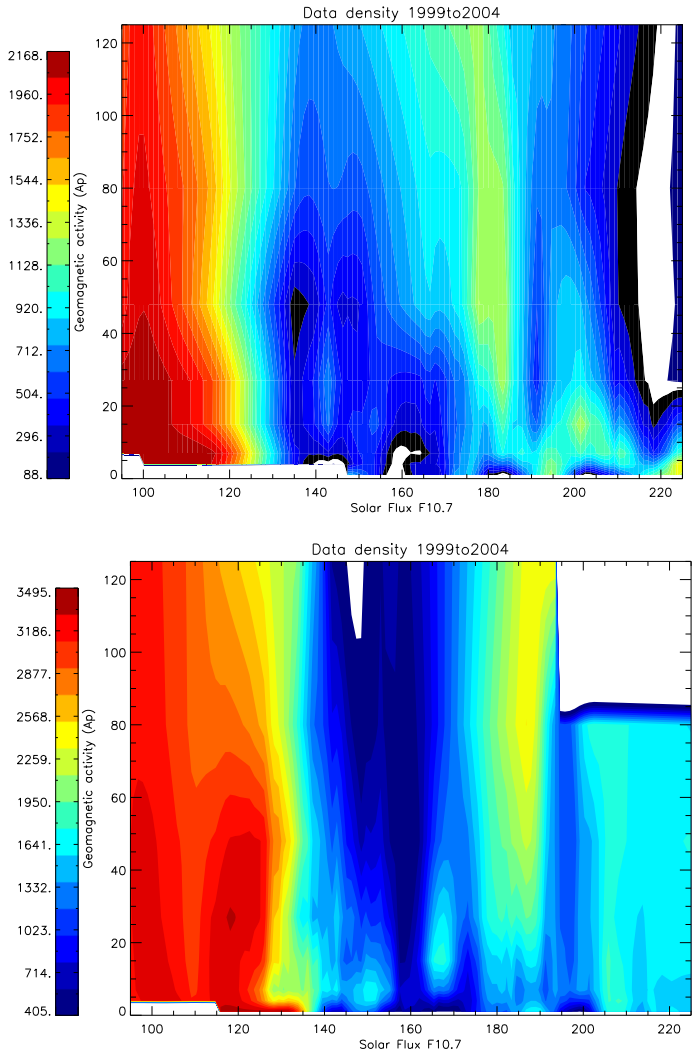


Figure 4.20: The plots here have the same format as figure 4.19. However, instead of plotting the contour of the neutral temperature at a particular set of Ap and F10.7, it shows the temperature data sample size (non-interpolated) that have been used for constructing figure 4.19. The results are from the solar maximum years 1999 to 2004.

data gap and interpolated results should be noted during interpretation.

Gaps in the MSIS data do exist because it requires contribution from satellite accelerometer, temperature from incoherent scatter radar, molecular oxygen number density from solar UV occultation etc. as described in section 2.6.1. Hence, the MSIS density does vary. Also MSIS data were acquired directly from the MSIS website, and not run locally in UCL, which explains some of the gaps.

However, the main reason for the variation/gap is that the data density plots reflect the number of temperature data points that fall within a certain pair of F10.7 and Ap

values. Hence, the density should not be uniform. We do not expect to have the same neutral temperature sample size for the moderate activity values of Ap and F10.7, as the very high or very low F10.7 and Ap value/conditions.

The general accepted view is that the low activity data are expected to be well modelled by MSIS, which are relatively smooth spatially and slow varying. The conclusion here shows that it is possible for MSIS to over estimate temperatures even under low activity conditions. It also points out the cause of the unusual MSIS T_n > FPI T_n demonstrated in the FPI temperature study in section 4.3.3.

It is also usually agreed and has been reported by a number of previous studies including Killeen et al. [1995] that model data are unable to recreate the high activity results, and therefore produce a lower T_n than the average measured temperature. This effect is also well demonstrated by the high activity region at $F_{10.7}$=180-230 and Kp=20-120. Overall in the solar maximum results, the MSIS data are unable to recreate both the high and low activity data. They underestimate temperature at the high activity region, but unexpectedly overestimate temperature at the low activity section.

4.3.6 The solar and geomagnetic conditions that may cause modelled and measured results to differ, solar minimum

Similarly, both the FPI and MSIS solar minimum temperature results are shown in figure 4.21, and the equivalent data densities are shown in figure 4.22. The $F_{10.7}$ range is 65-95 and Ap range is 0 to 80 only for the extended solar minimum period. Again, comparing with the average temperature results and distributions shown in figures 4.15 and 4.16; the main conclusion from section 4.3.3 is that the average MSIS temperature is lower than the FPI measured temperature at solar minimum. The contour plots of temperature and density of $F_{10.7}$ vs Ap (figures 4.21 and 4.22) demonstrate MSIS underestimates neutral temperature at both high and low activity regions. This is different from the solar maximum results. MSIS is an empirical model that applies years of observations from both ground and space and therefore may not be able to accurately represent the unusual extended solar minimum condition.

The average solar minimum temperatures shown are also similar to the Killeen et al. [1995] results from the previous solar cycle. Therefore the local polar cap, F-region winter temperatures were not lowered significantly during the extended solar minimum, comparing to the previous solar cycle. This result appears to contradict the unusually low neutral

mass density for this period reported by Emmert et al. [2010], which is associated with low thermospheric temperature. However, results from Emmert et al. [2010] are global average results. The FPI temperatures here may suggest that the high latitude regions are affected differently from the global average under the extremely quiet conditions, possibly because of the influence of heating due to local geomagnetic activity.

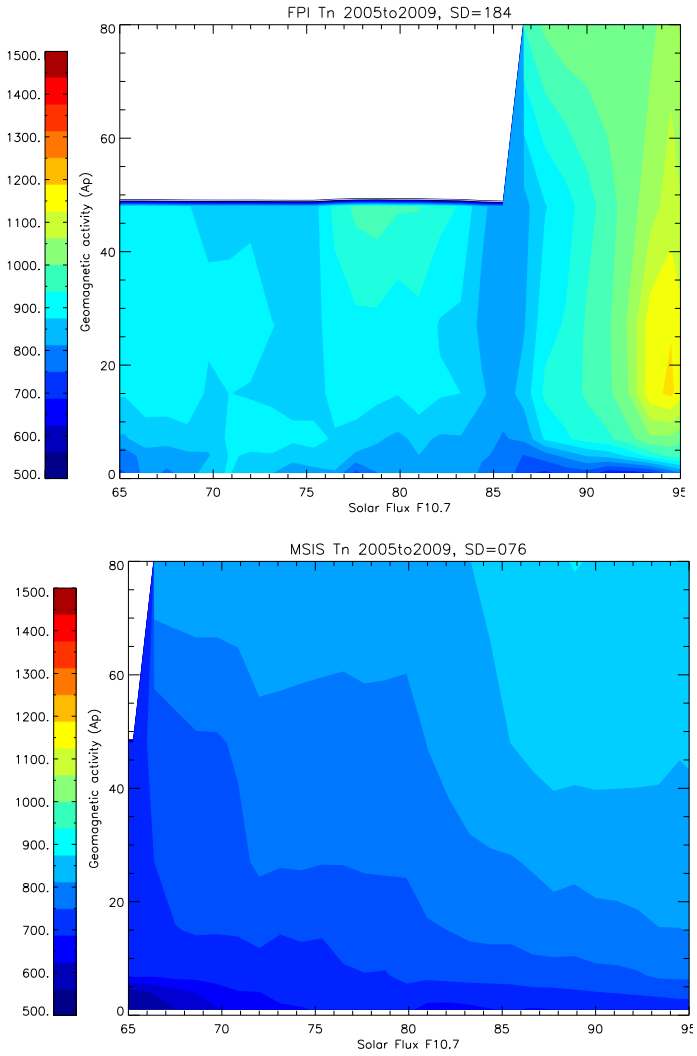


Figure 4.21: These are the hourly Svalbard FPI T_n measurements vs. Ap and $F_{10.7}$ on the top, and the equivalent MSIS T_n measurements on the bottom, for solar minimum years 2005 to 2009. The standard deviation (SD) for all the data points included in the plot are provided.

4.3.6.1 Solar cycle trend summary

- For temperature trend investigations, daily-averaged results are preferred because there is no bias towards any time in the daily diurnal variation. It also has the

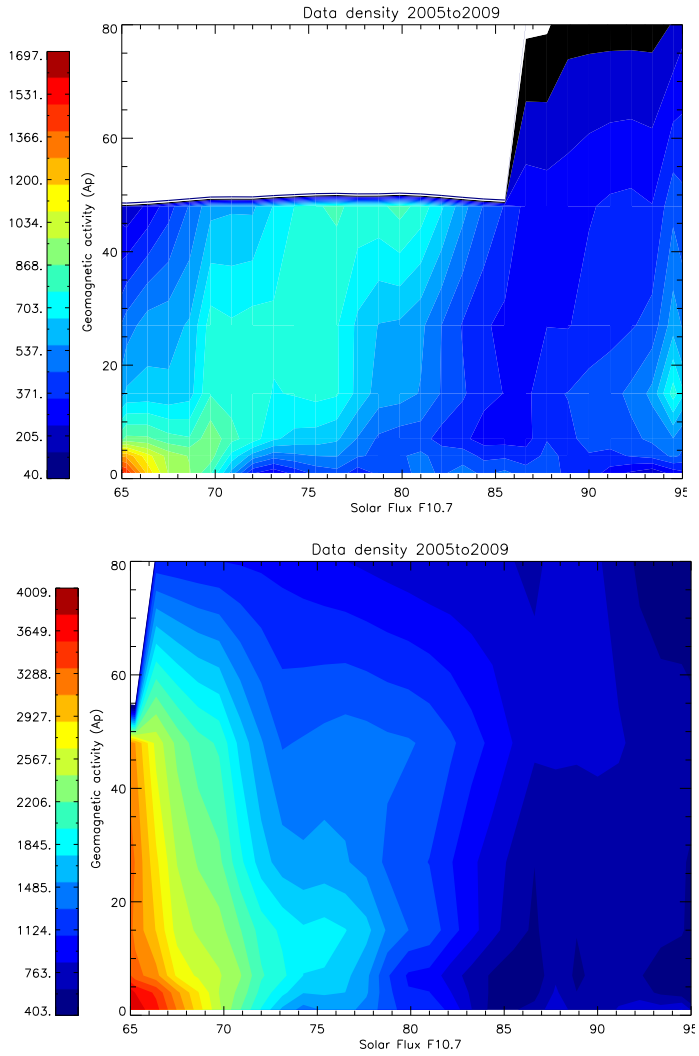


Figure 4.22: The plots here have the same format as figure 4.21. However, instead of plotting the contour of the neutral temperature at a particular set of Ap and F10.7, it shows the temperature data sample size (non-interpolated) that have been used for constructing figure 4.21. The results are from the solar minimum years 2005 to 2009.

advantage of minimising the effects of clouds. For Svalbard, full 24-hours daily-averaged data can only be obtained during November, December and January, and these are the months chosen for the yearly/season average T_n calculations. Biasing towards a particular time period could introduce a ~50K error in the final averaged-T_n results.

- For solar maximum between 2001 and 2004, the FPI-measured T_n shows an extended, but unexpectedly lower than MSIS-modelled T_n peak. The measurement is

lower than modelled by more than 200K. The correlation for both T_n vs. $F_{10.7}$ and T_n vs. Kp for daily average data are low.

- Another possible small contribution to the lower measured T_n than modelled T_n at solar maximum is the wider red-line emission peak. Hence the FPI measurements are detecting more of the lower T_n in the T_n vertical profile. It is expected to cause the FPI T_n to be lower by 25-50K. The shape of the red-line emission peak is an important factor for a long-term FPI temperature trend study. Further testing then shows that MSIS unexpectedly produces higher than FPI-measured T_n for the low Kp and $F_{10.7}$ conditions at solar maximum. The lower Kp and $F_{10.7}$ conditions also contain the highest number of data points, and dominates the data set used to calculate the overall average. Hence resulting in MSIS $T_n >$ FPI T_n for solar maximum.

- For solar minimum between 2008 and 2009, the average FPI-measured T_n results are higher than the modelled results by ∼60K. The actual difference is expected to be ∼100K if the lower red-line emission peak is taken into consideration. The average measured T_n is ∼750K at this extended solar minimum, which is similar to the average values from the previous solar minimum [Killeen et al., 1995]. This appears to contradict the reported low global neutral density and hence temperature, and suggests local geomagnetic influence is more important than expected for the polar cap region. Overall, the correlation for both T_n vs. $F_{10.7}$ and T_n vs. Kp, and both daily and hourly average data, are low. Further testing then shows that MSIS produces lower than FPI-measured T_n for all conditions at solar minimum.

- The best match between model and measurement is found between 2006 and 2007 at the beginning of the solar minimum. Overall for this solar cycle, the average measured winter T_n has an unexpectedly smaller variation than modelled results over the solar cycle, where the measured T_n is lower than modelled on average at solar maximum; and measured T_n is higher than modelled on average at the extended solar minimum.

- The comparison of different proxies, $F_{10.7}$, $S_{10.7}$, Kp and AE show no significant signatures in the correlation results. This demonstrates that for the polar cap, average winter F-region temperature, the proxies are unlikely to be the cause of the model-measurement discrepancies.

4.3.7 Diurnal UT trend

4.3.7.1 UT-binned data

The FPI/SCANDI temperature data binned according to UT from December of each observing season are averaged, and shown in figure 4.23. Also, MSIS data are generated for each hour for every day in December for each year.

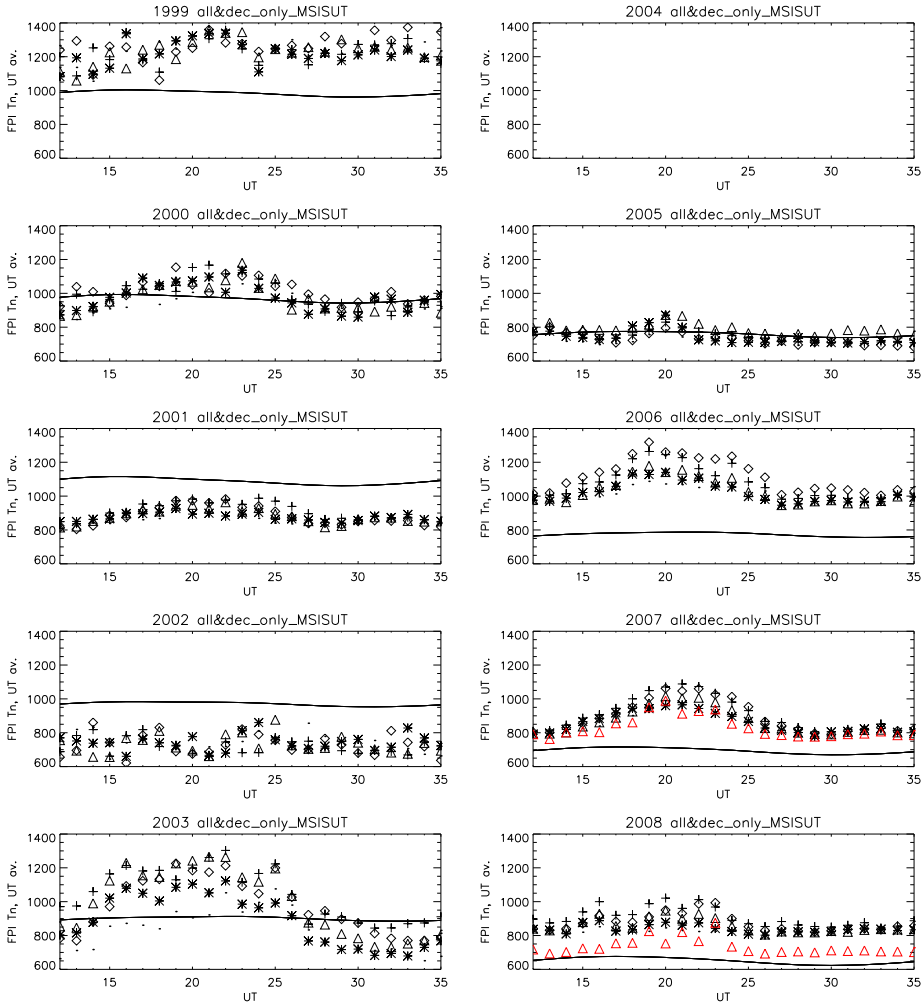


Figure 4.23: These are the December T_n data averaged according to universal time (UT) from years 1999 to 2008. The data plotted with the various symbols are the measured data, with the FPI-data in black and SCANDI data in red. The different symbols represent measurements from different line-of-sight directions. The lines are the modelled results from MSIS. Each figure represent data from December of each year, for years 1999 to 2003 from top to bottom in the left column, and from 2004 to 2008 in the right column. There is no 2004 data.

The measurements follow a general diurnal variation, which match with previous studies by e.g. Killeen et al. [1995] and McCormac et al. [1987]. The diurnal variation on

average gives a T_n variation of ∼200-300K from peak to trough, but MSIS shows a much smaller variation of less ∼50K in all years. The diurnal T_n peak is found between 20 and 22UT (23-25MLT), and the minimum is found between 29 and 30UT i.e. 4-5UT (7-8MLT). A smaller additional peak with a T_n elevation of 50-100K is found at 32-34UT i.e. 8-10UT (11-13MLT), which matches with the time of cusp-crossing. The neutral perspective of the cusp is investigated fully in the next chapter.

4.3.7.2 CMAT2 model comparison

The source of the diurnal pattern has been discussed in detail by previous studies including Killeen et al. [1995] and McCormac et al. [1987]. They both consider a number of heating and cooling terms listed in section 2.2.4, which are analysed by the thermospheric models of their choice. Since advection is one of the major factors affecting the neutral temperatures, both authors also consider the heating history of the air parcels that are detected by their instruments, by modelling the trajectory of the air parcels starting from 24 hours prior to detection [McCormac et al., 1987]. It shows that the air parcels do retain some of the heat. The final temperature depends on whether an air parcel has passed over e.g. the cusp, the day-side, the different parts of the circulation pattern, before being measured.

Over the solar cycle, the UT variation of the diurnal pattern remains similar, apart from the magnitudes of the peaks and troughs [Killeen et al., 1995]. Therefore only the December 2007 modelled results are shown in figure 4.24 for FPI-T_n and modelled-T_n comparison. CMAT2 [Harris, 2001] is the model chosen here. It should be noted that CMAT2 is a time-dependent global 3D model, and its high latitude spatial resolution is therefore not comparable to e.g. 2D standard ionospheric conductivity models that focus on height and particle density dimensions only. CMAT2 has a longitude resolution of 18 degrees, latitude resolution of 2 degrees and time resolution of 1 hour.

The most important heating terms from CMAT2 are shown in figure 4.24 for FPI comparisons. Over the whole 24 hours, the major terms are Joule heating, horizontal advection and particle precipitation heating, which cover all precipitation events including those from aurora and cusp. The most significant cooling terms are horizontal advection and vertical conduction. These results are similar to previous conclusions [Killeen et al., 1995, McCormac et al., 1987] where the diurnal pattern is mainly driven by the horizontal advection heating term, while Joule heating and particle precipitation can also become important because of high latitudes.

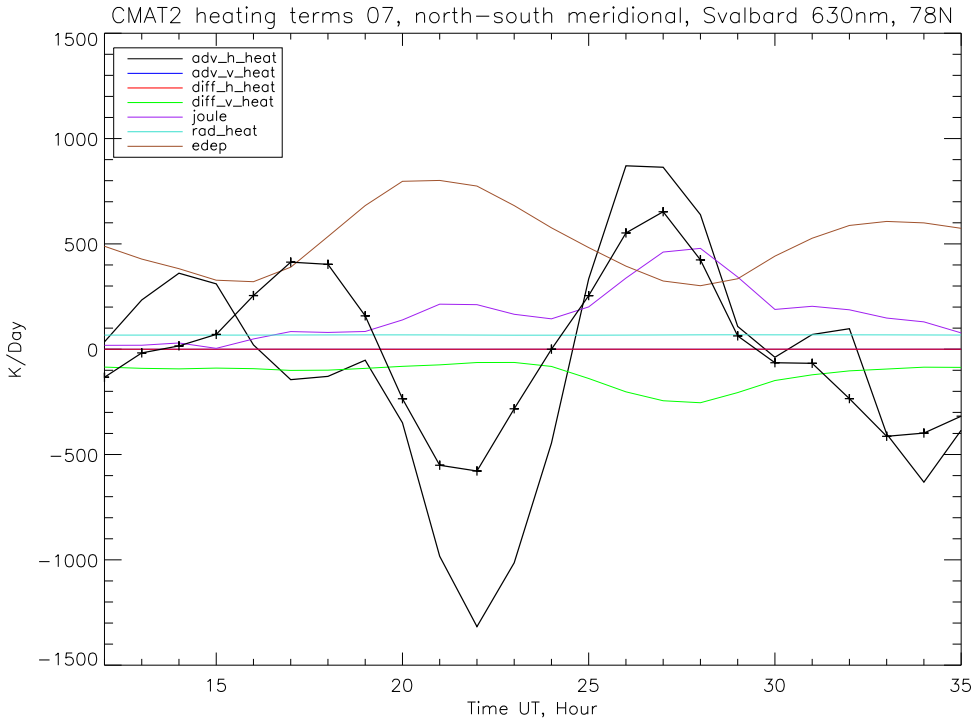


Figure 4.24: This figure shows the diurnal variation of heating terms output by CMAT2, and are the modelled results representing December 2007. The data are appropriate for the centre of the SCANDI FOV. The total heating/cooling is marked by the line with the plus signs. The individual heating terms that are considered by the model are shown in black for horizontal advection, blue for vertical advection, red for horizontal conduction, green for vertical heat conduction, purple for Joule heating, turquoise for radiation, and brown for particle precipitations, which include all precipitation events i.e. auroral and cusp-effect. The black line with crosses represents the overall relative T_n variation. The relative T_n variation is plotted here, matching with the other parameters in this plot. Note blue and red graphs overlap each other. The absolute temperature variation can be found in figure 4.26.

Comparing the CMAT2 T_n (figure 4.24) to the general measured diurnal T_n pattern (figure 4.23), the similarity is that both model and measurement give two peaks, but the relative peak magnitudes are significantly different, while the timings of the peaks are approximately the same. An extra study is performed by utilising the high spatial resolution available from SCANDI. The average meridional temperatures from SCANDI are taken from the zones illustrated in figure 4.25. These are compared with the meridional variation in the CMAT2 modelled results, and are shown in figure 4.26.

Consider the first peak, which occurs at similar times at 20UT (23MLT) for measurement and model, but the measured peak (figure 4.26, top plot) has a wider temporal

width that extends between 18 and 22UT (21-25MLT). The measured first peak also has a temperature of 800-900K, while the modelled first peak (figure 4.26, bottom plot) is lower at 700K. The measured peak centres at higher latitude above 80 degrees, but the modelled peak is wider along the meridian, and extends to lower latitudes. The CMAT2 auroral input is modelled and changed according to the precipitation efficiency curve in response to the Kp value variation (section 2.2.4). CMAT2 uses three hour resolution Kp data but it can use any time step and output data at any time resolution. Higher spatial resolution Kp values are also available and are used in CMAT2 for the ESA ATMOP satellite drag study.

Comparing the SCANDI temperature results (figure 4.26, top plot) to the heating terms shown in figure 4.24, the main contributions to the first temperature peak between 17-23 UT is electron precipitation heating, which includes aurora and other sources of particle precipitations. The precipitation heating remains similar for all modelled latitudes over SCANDI FOV. Note that the standard spatial resolution of CMAT2 is 2 degrees latitude, which represents five data points in latitude within the SCANDI FOV. The CMAT2 horizontal advection is a maximum in the northern most zone of the SCANDI FOV (not shown), and is a cooling term at the lower latitudes for the period containing first temperature peak observed in the central SCANDI zone, as shown in the plot. The high rate of heating by precipitation is likely to be caused by auroral precipitation from the contracted auroral oval at this solar minimum. The change from heating to cooling from

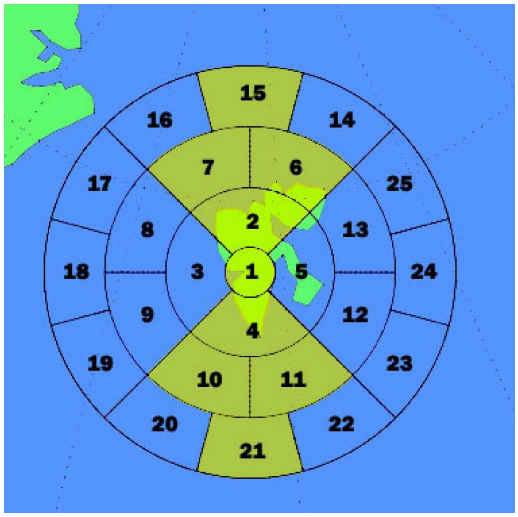


Figure 4.25: This figure shows the 25 SCANDI observation zones, where the meridional zones are highlighted in yellow. The meridional temperatures are those from zones: 15, averaged between 7 and 6, 2, 1, 4, averaged between 10 and 11, and finally 21.

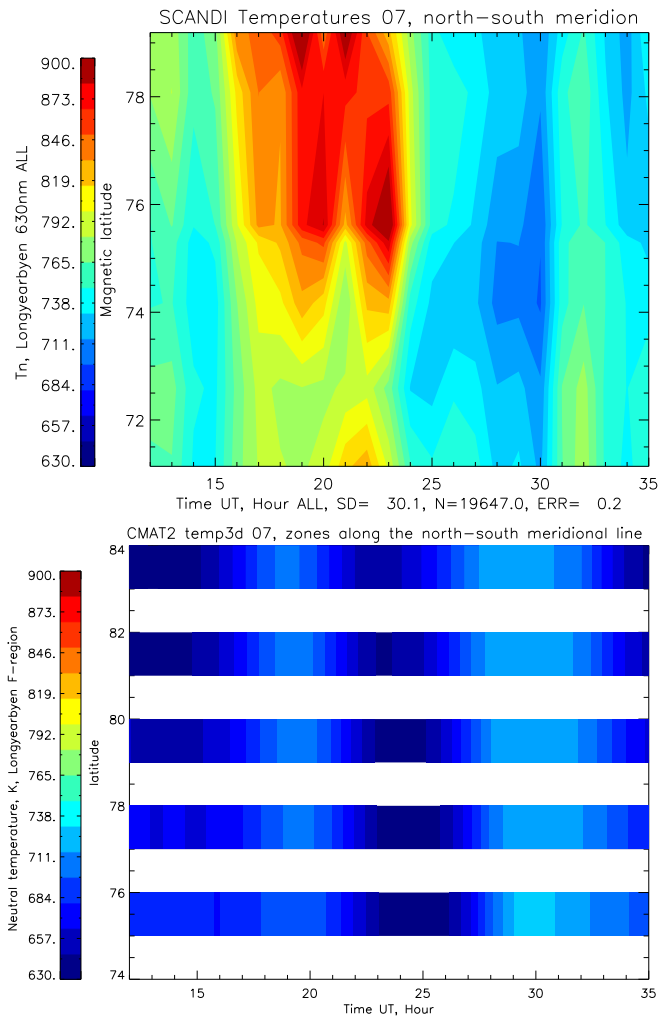


Figure 4.26: These figures are the December 2007 averaged, meridional diurnal T_n variation. The plot on the top is the SCANDI measured results, and the plot on the bottom is the CMAT2 modelled results. The y-axis of each plot is the latitude, which covers the meridional SCANDI FOV shown in figure 4.25, and the x-axis is UT. The T_n plotting range is the same for the plots. The CMAT2 data are produced over 1 degree latitude area, for every 2 degrees latitude, which is why the CMAT2 data are presented as 1 degree bands.

high to low latitudes for horizontal advection is expected. However, the total magnitude of heating from the model is unable to replicate the magnitude of the measured T_n. CMAT2 is a global model and the results here demonstrate the importance of the development of a higher spatial resolution nested grid for the auroral region.

The second T_n peak occurs between 30 and 35 UT i.e. 6-11UT (9-14MLT) for the measurements, while the modelled second peak is temporally wider at 28-35UT i.e. 4-11UT (7-14MLT) (figure 4.26). The temperatures are similar between measurement and

model, which are ∼760-780K and ∼710-730K respectively. The measured peak centres around lower latitudes within the observing area, while the modelled peak covers the entire meridional FOV.

Again, comparing the SCANDI second T_n peak results (figure 4.26, top plot) to the different heating terms in figure 4.24, the main heating terms for the second peak are horizontal advection, Joule heating and particle precipitation heating. Joule heating is directly related to the ion-neutral collision and the difference between neutral wind and ion velocity i.e. a balance between solar heating pressure which drives a general anti-sunward neutral wind flow, and the general ion circulation, typically a 'twin-cell' pattern which can change drastically when the IMF changes e.g. Weimer [2005]. Hence, a high Joule heating rate is generally expected near dawn and dusk at lower latitudes because of the sunward flow of ions, which is against the general anti-sunward neutral flow due to solar pressure. This is a simple view appropriate for the average neutral conditions that are considered here, which are at resolutions of hours. For a polar cap latitude, the highest Joule heating regions could be closer to the ∼12MLT time period here, as shown in figure 4.27 (courtesy Weimer [2005]). These nine plots indicate precisely the temperature variations that are demonstrated by both SCANDI and CMAT2 data, and support that the highest temperature is found at around 12 MLT over the Svalbard magnetic latitude of 74 degrees due to Joule heating.

The common cause of the midnight sector peak not being noted is because the overall winds are too large in most thermospheric models. The reason is most existing models do not allow for variation in the electric field and assume steady state wind. Hence the wind is too high and cooling term becomes too large. It is possible that the mid-night sector peak is not yet modelled by CMAT2 with a magnitude that matches with observation. There should also be consideration that this is a data set from a very low activity period with a contracted auroral oval, hence the heating from direct particle precipitation is stronger.

In terms of the elevated precipitation heating, the second T_n peak matches with the position of the cusp as Svalbard passes underneath. The cusp does not cover the entire SCANDI meridional FOV. SCANDI should be able to spatially resolve the cusp region with both its wind and temperature data. This is presented in the next chapter.

4.3.7.3 UT variation Summary

- This section presents the first study of the CMAT2 model results with FPI measure-

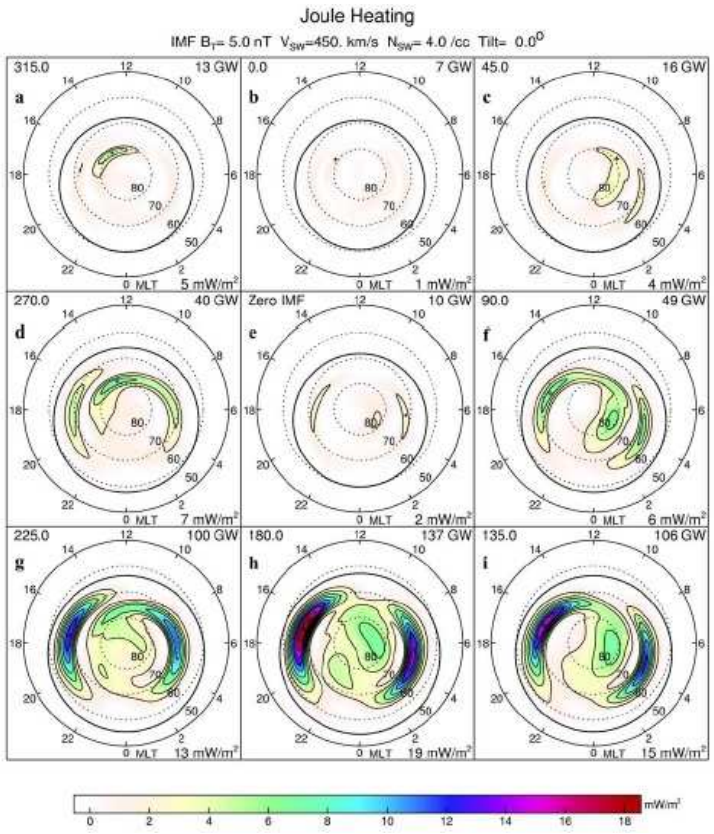


Figure 4.27: The plots show the Joule heating rate in the ionosphere in the Northern Hemisphere at an altitude of 110km (E-region), courtesy Weimer [2005]. The results shown in this thesis section are from a higher altitude of 240km (F-region. Each plot shows the Joule heating predicted by Weimer [2005]'s model at different IMF tilt angle, printed at the top-left corner of each plot. The total integrated heating is printed at the top-right corner.

ments. It includes a discussion of the possible CMAT2 heating sources that may have caused the averaged diurnal pattern seen in the FPI measurements. The high spatial resolution SCANDI meridional temperatures are also compared against the CMAT2 results.

- For the UT-binned data, the measured T_n remains lower than MSIS-modelled T_n at solar maximum, and the opposite for solar minimum. This result is similar to the daily-averaged solar-cycle trend conclusion from the previous section.

- The measured diurnal T_n variation averaged around 200-300K from trough to peak, but the MSIS model only shows variations of about 50K.

- The first diurnal peak is found at around 21-25MLT (18-22UT), and investigation of the CMAT2 heating sources suggest the main heating contribution is particle

precipitation heating. The measured T_n peak is higher than the CMAT2 modelled peak by about 150K. This result demonstrates the need for higher spatial resolution for the polar region for circulation global models.

- A second smaller peak is found at around 9-14MLT (30-35UT), and the measured and CMAT2 modelled peaks are similar at about 750K. CMAT2 output of different heating terms suggest the main contributions to this peak are horizontal advection, Joule heating and particle precipitation heating.

Chapter 5

The cusp and the neutral atmosphere in the F-region

5.1 Introduction and scientific goal

There are still uncertainties involved in the neutral density variations relating to the cusp. For example, measurements at lower altitudes of 180-350km show density enhancements around the cusp and hence a relative density decrease within the cusp [Clemmons et al., 2008]. Other measurements from e.g. CHAMP [Lühr et al., 2004, Schlegel et al., 2005] are taken from a higher altitude of above 400km. Both papers demonstrate a cusp density enhancement by up to a factor of two on their average data.

Furthermore, as summarised by Carlson [2007], the F-region plays an important role in the density changes observed at higher altitude. The soft cusp particle energy fluxes are typically deposited at about 140 km, largely in the F-F1 region, in contrast to hard particle precipitation penetrating to the E-region. The E-region has larger neutral density and smaller neutral scale height (about 10 vs. several 10s of km). The atmosphere also reacts more dramatically in the F-region, mostly through adiabatic expansion against gravity; where the heated cell then expand upwards. Downward Poynting flux energy is dissipated by the Pedersen (parallel to E) component. At this height, Pederson conductivities exceed Hall, and hence it is also the altitude where downward Poynting flux is largely dissipated. The Poynting flux is the electromagnetic power per unit area generated by the solar wind-magnetosphere dynamo. Equation 2.28 describes how the Poynting flux is redistributed in the thermosphere as Joule heating rate plus rate of work done in the acceleration of the neutral gas. These elements help to amplify the overhead cusp response from the F-region

altitudes.

Direct measurements of neutral winds and temperatures can give further insight into how the neutral atmosphere is involved in the density enhancements in the cusp. The neutral wind dynamo contributes to E-field as part of Joule heating. The vertical winds act as evidence of up-welling, and enhancement in neutral temperature as evidence for heating.

5.2 The suitability of FPI data

Previous ground-based studies of the cusp focus on ionospheric measurements, and consider changes in electric fields and charged particle fluxes. This is mostly because neutral measurements at the lower thermosphere are sparse and are technically challenging.

The results here present the first measured neutral parameters from a FPI located in Longyearbyen, Svalbard. In winter, Longyearbyen is under 24 hours of darkness, which enables optical observations of the cusp, only restricted by sky conditions such as cloudiness. The study here presents statistical averages of FPI parameters of F-region neutral temperature, line-of-sight wind and red-line intensity 630.0 nm, separated into samples according to solar maximum/minimum and geomagnetic quiet/active conditions. The directly observed neutral results help to identify the typical characteristics of neutral temperature and winds relevant to the cusp region, which have not been presented before.

Three case studies are also presented in support of the statistical results. They demonstrate well that the cusp is indeed within the FPI/SCANDI field-of-view (FOV), with supporting evidence from the co-located EISCAT radar ionospheric results and Meridional-Scanning Photometer (MSP) optical measurements along the geomagnetic meridian.

5.3 Data and instrumentation

The statistical FPI/SCANDI data are sampled as follow:

- Solar maximum years (1999-2004), quiet geomagnetic conditions
- Solar maximum years (1999-2004), moderate/active geomagnetic conditions
- Solar minimum years (2005-2009), quiet geomagnetic conditions
- Solar minimum years (2005-2009), moderate/active geomagnetic conditions

The level of geomagnetic activity is determined by using Kp values. It is then roughly assumed that quiet geomagnetic samples correspond, in general, to weak or northward IMF Bz, and to more complicated cusp dynamics. The moderate/active geomagnetic samples correspond to stronger and southern IMF Bz.

Three case studies are presented to support the results of the statistical data. Two of the case studies represent southward IMF Bz conditions, where cusp dynamics are well understood. FPI data are used throughout and the cusp was within the FOV. The remaining case study covers a strong IMF By night under generally quiet conditions at solar minimum, with more complex cusp up-welling mechanism. The higher spatial resolution SCANDI data are used. The relationship between the IMF and the cusp is well demonstrated in figure 2.18 and described further in section 2.5.3. They show the effect of changing IMF tilt angle on the electric potential in the northern hemisphere, and hence the location of the cusp.

The three case studies are:

1. Steady southward IMF Bz on 17^{th} December, 2001. The cusp was detected towards the southern side of FPI's field-of-view (FOV).

2. Steady southward IMF Bz on 18^{th} December, 2001. The cusp was detected towards zenith/more northern side of FPI's field-of-view.

3. Dominating IMF By, 10th January, 2008. The cusp was detected toward northern side of SCANDI's field-of-view. SCANDI high spatial resolution is used.

5.4 Case studies 1 and 2 - settings

The first two case studies were two very similar experiments from two consecutive nights on 17^{th} and 18^{th} December, 2001, in Svalbard. The instrumentation and their set-up were also the same. The ground-based instruments involved were the FPI, EISCAT ESR radar and the MSP photometer/Keogram. The FPI was observing red-line emission which peaks in the F-region. The EISCAT and MSP/Keogram data from the 17^{th}, between 06:50-07:35 UT were also used in an ionospheric patches study by Carlson et al. [2006], where the cusp was clearly identified in both the EISCAT and MSP data/Keogram. The ionospheric mechanisms over this time period were also well documented and understood in the paper, but neutral measurements were not considered. Hence, the neutral atmospheric data from the FPI data were able to provide a valuable neutral perspective to this time period. This

case study can also provide important information for interpreting the statistical FPI data over the recent solar cycle.

The observing locations of the instruments projected at 240km altitude are shown in figures 5.1 and 5.2. They are presented in geographic and geomagnetic co-ordinates respectively in the figures.

Using case study 1 from 17^{th} December 2001 and figure 5.1 as an example: the FPI was observing geographic North, East, South, West, Zenith, Southeast (SE) and Southwest (SW) as labelled in geomagnetic co-ordinates in figure 5.2. Each FPI observing-position required 1 minute integration, and the FPI observing-mirror rotated through each position consecutively. A complete mirror rotation took 15 minutes. The MSP was 'windshield-wiping' between 30 degrees elevation north/south via the zenith, along the magnetic meridian.

For the cusp case studies here, the general EISCAT radar observing-points at 240km were positions 1 to 4 between 6 and 7 UT, positions 5 to 8 between 7 and 10 UT, and positions 9 to 12 between 10 and 11 UT. Each scan required 2 minutes for the EISCAT meridian scan period at positions 5 to 8. Two different radar scan modes were deployed for the 17^{th} December 2001, case study 1 and 18^{th} December 2001, case study 2:

- For case study 1, there was a 'windshield-wiper' mode to the right of the line of magnetic meridian, between 6-7 UT. This was followed by a four-point scan along the magnetic meridian, which matched the MSP scan at 7-10 UT. Then from 10-11 UT, there was a 'windshield-wiper' scan to the left of the magnetic meridian. The mode described here follow the general cusp ESR settings.

- For case study 2, the sequence of east side 'windshield-wiping', followed by overhead and then west side, was repeated for a different period of time at 6:30-8 UT, 8-9:30 UT and 9:30-11 UT respectively. Hence, the sequence of positioning remains the same as the general ESR cusp setting, but the timings are different. The aim was to keep the cusp in view as the observing site passed underneath.

5.5 A note on graph interpretation

When interpreting and comparing the Keogram/MSP, EISCAT, and FPI data in, for example, (case study 1) figures 5.3, 5.4, 5.5, 5.7, and 5.8, it is important to note their

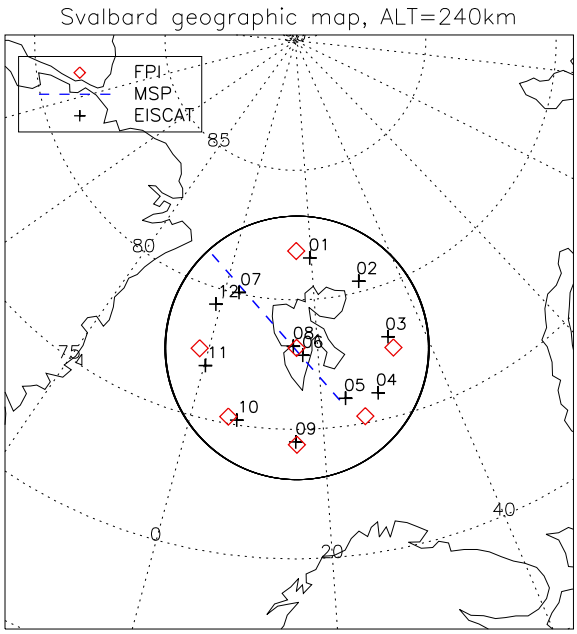


Figure 5.1: This figure shows the FPI, MSP/Keogram and EISCAT observing locations, projected at 240km altitude, above Longyearbyen, Svalbard. The red diamonds represent the FPI, the black crosses are EISCAT points and the blue dashed line is the MSP scanning direction. This is plotted in geographic latitude and longitude co-ordinates.

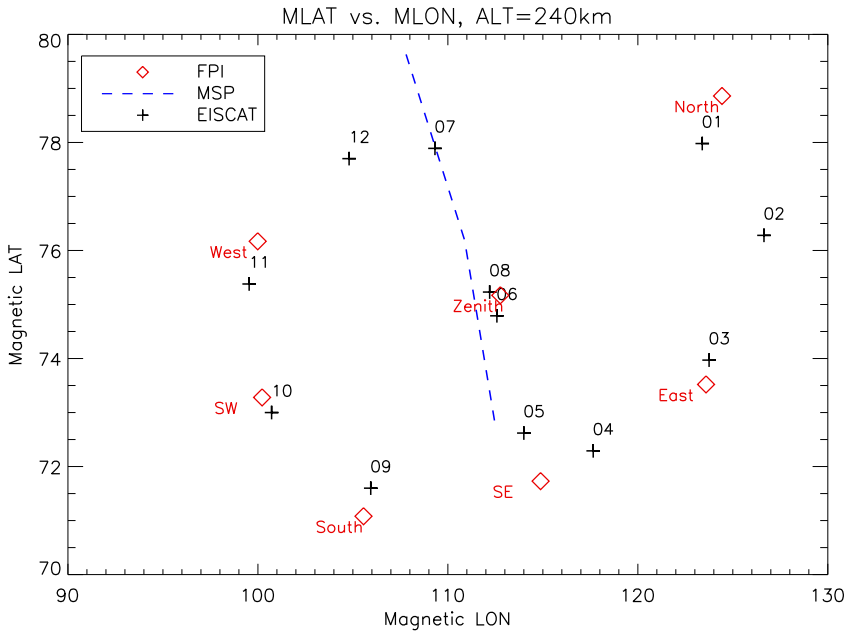


Figure 5.2: This is similar to figure 5.1, but the plotting co-ordinates are in geomagnetic longitude and latitude.

longitudinal variations and their relative positions using the accompanying map figures 5.1 and 5.2.

For example, the Zenith FPI position should be compared with EISCAT position 08 and 06, FPI east with EISCAT position 03 etc. The MSP/keogram observation is along the magnetic North-South and should be compared with the geographic North-East, Zenith and South-East. Additional plots where only EISCAT data along the geographic north-zenith-south are also included (figures 5.6 and 5.12). FPI data matching with the geographic north-zenith-south only are also high-lighted in red.

Please note plotting the LOS Vi, as well as the LOS FPI wind from all available directions in the same plot makes it possible to estimate the overall Vi and wind direction. All FPI and EISCAT data are plotted with magnetic latitude as their vertical axis for consistency, and to follow previous cusp studies e.g. Carlson [2007].

5.6 Case study 1 - 17^{th} December 2001

5.6.1 Keogram

The Keogram data from case studies 1 are shown in figure 5.3. The cusp can be clearly identified within the observing period between 6 and 11 UT. Soft cusp precipitations can be identified when the red to green-line intensity ratio is high. The cusp is seen mostly towards the southern side of the Keogram's field-of-view (FOV).

5.6.2 EISCAT

The EISCAT data from case study 1 are shown in figure 5.4. The F-region EISCAT electron density (Ne), electron temperature (Te), ion temperature (Ti) and line-of-sight ion velocity (LOS Vi) are displayed in the figure. The most easily identified feature is the increased F-region Ne, which match well with the 630 nm intensity variations shown by the Keogram in figure 5.3, relating to the cusp. The electron densities shown are smaller or equal to $10^{11.5}$ m^3 for magnetic latitude above 72.5 degree. This is typical of the dark polar F-region. The electron and ion temperatures reach 1500K for the cusp regions with increased electron density. The highest temperatures are found towards the south, also matching with the cusp location. Case study 1 shows a general flow from the northwest at greater than 500 m/s for between 7 and 10 UT. This is consistent with a By positive and Bz negative (southern) IMF condition. The By positive and Bz negative (southward)

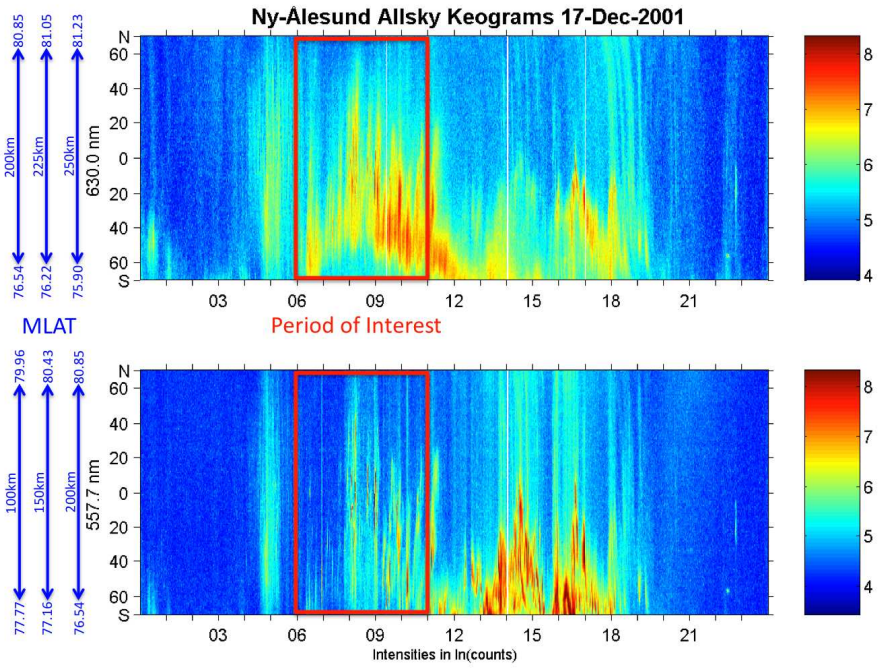


Figure 5.3: The top and bottom plots are the red (630.0 nm) and green-line (557.7 nm) Keogram intensity data respectively for 17^{th} December, 2001. In addition to the zenith angle as the y-axis in black, the blue axis demonstrates the projected range in kilometres and magnetic latitude at 3 altitudes around the expected peak emission height.

IMF condition gives a simpler twin cell pattern, where the effect of ion drag on the neutral atmosphere can be more easily predicted, as shown in figure 2.18.

5.6.3 FPI

The data shown so far are some of the most commonly observed parameters for cusp studies. They concentrate on photometer and ionospheric parameter measurements. This section provides new and additional simultaneous observations of the neutral F-region by the FPI, under a relatively well-understood cusp in terms of the ionosphere, under IMF Bz southward condition [Carlson et al., 2006].

5.6.3.1 FPI intensity

The FPI intensities in figure 5.7 can be compared to the MSP data in figure 5.3. The look angle of the FPI covers 67 degrees from zenith. The different azimuth positions therefore provide measurements corresponding to a range of latitudes within the limits of ±67 degrees look angle along the meridian. This is similar to the range of latitudes viewed

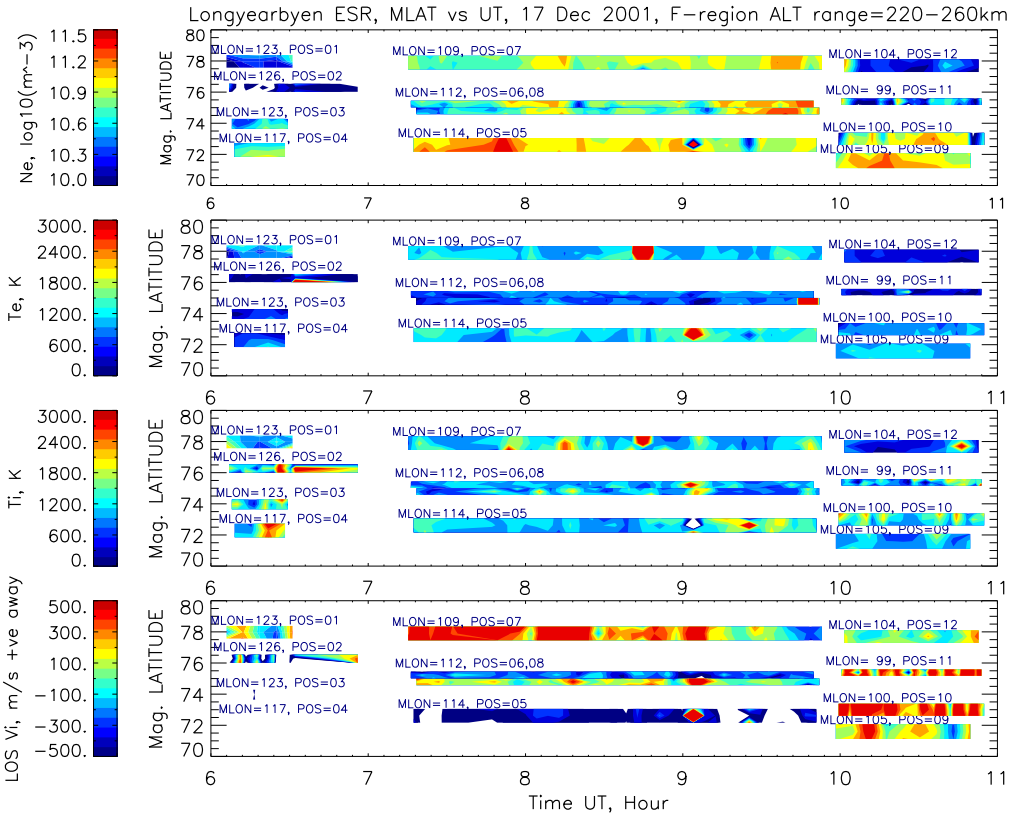


Figure 5.4: This is the EISCAT data from 17th December 2001 plotted in magnetic latitude vs. time in UT. The top plot is electron density, follow by electron temperature, ion temperature and line-of-sight ion velocity. Only data from the altitude range 220-260km are shown, representing the F-region. The colour bands' projected positions are as those numbered in figures 5.1 and 5.2. Note that each position's magnetic longitude is different. This represents a radar pointing sequence consisting of 'windshield wiping' to the east of the magnetic meridian, followed by overhead and finally to the west.

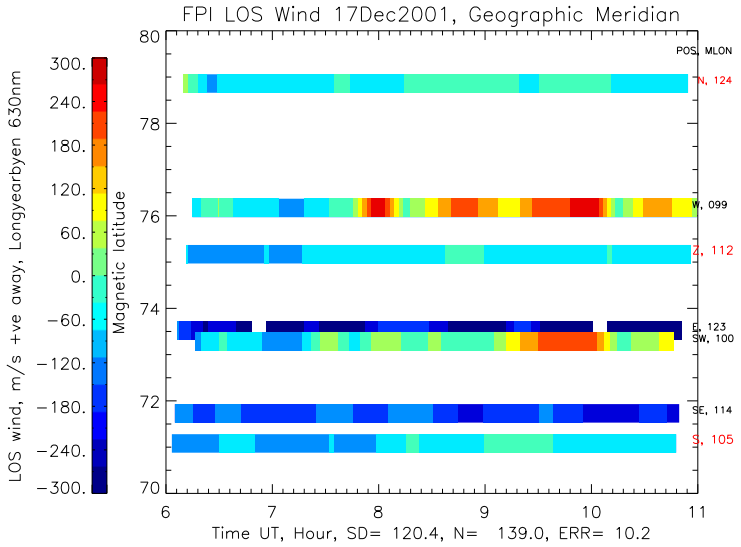


Figure 5.5: This is the FPI LOS neutral wind from 17th December, 2001, shown in the same format as figure 5.7.

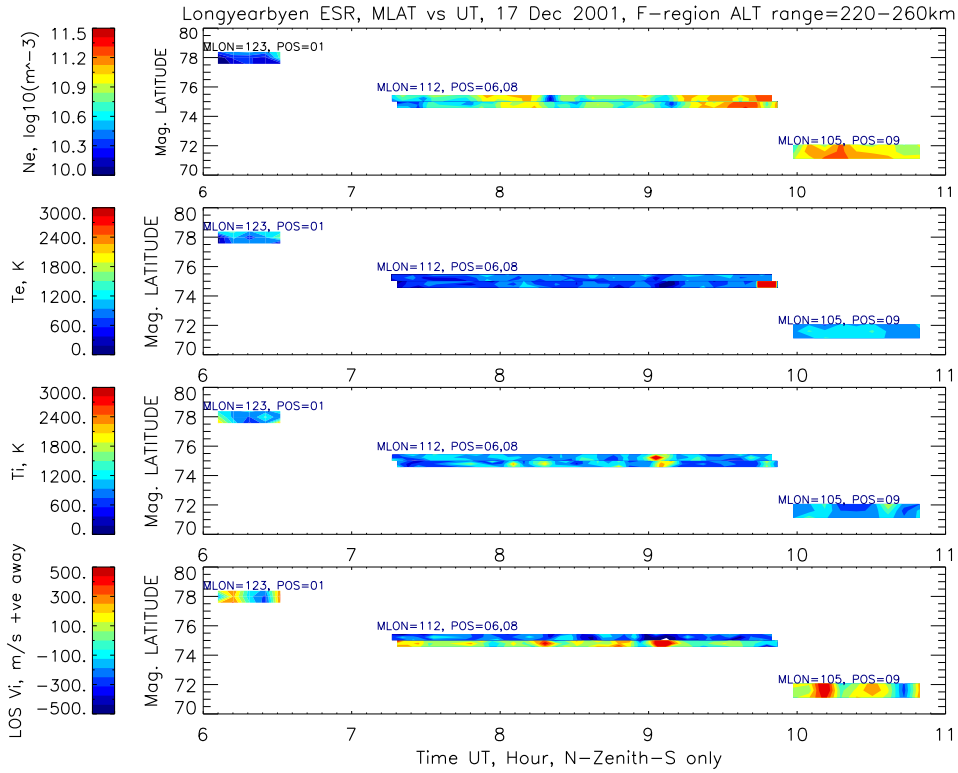


Figure 5.6: This figure is similar to figure 5.4, but only the geographic north-zenith-south directions are shown.

by the Keogram, so both figures 5.7 and 5.3 may be compared directly. However, only the FPI zenith and SE positions are along the Keogram. The detailed relative observing positions of FPI and MSP at 240km altitude can be found in figure 5.1. Considering the above, as well as that the FPI was observing at lower time resolution; the FPI and Keogram intensities are reasonably well matched and both show a clear red-line intensity enhancement caused by cusp precipitation. It can also be seen that the highest cusp red-line intensity is found towards the southern side for case study 1.

5.6.3.2 FPI temperature

The FPI temperatures in figure 5.8 can be compared to the FPI intensities in figure 5.7. The most distinctive feature is that the regions of highest neutral temperature enhancements do not overlap with the regions of highest intensity enhancements. For case study 1, the temperature is most increased between 6 to 8 UT, while the highest intensities follow at 9 to 11 UT, approximately divided between earlier and later UTs along the zonal direction. The case study indicates that the temperature and intensity enhancements in

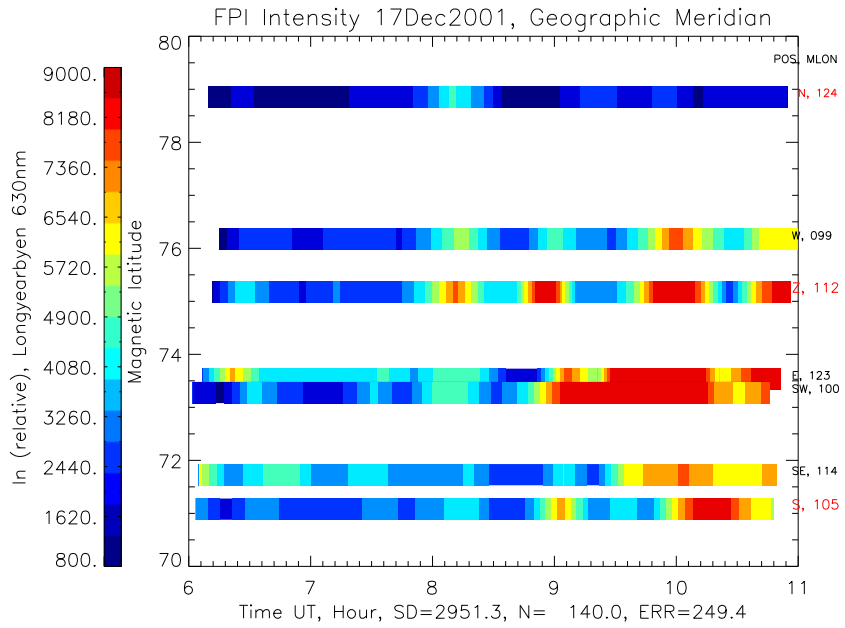


Figure 5.7: This is the FPI 630.0nm intensity plot for 17^{th} December 2001. Each colour 'band' represents data taken from positions north (N, highlight in red), south (S, highlight in red), east (E), west (W), zenith (Z, highlight in red), southwest (SW) and southeast (SE). Their magnetic longitude and position is labelled on the right side of the plot. Their projected position at 240km can be found in figures 5.1 and 5.2. The standard deviation (SD), number of data points within the plot (N) and error (ERR) calculated as SD/sqrt(N) is included.

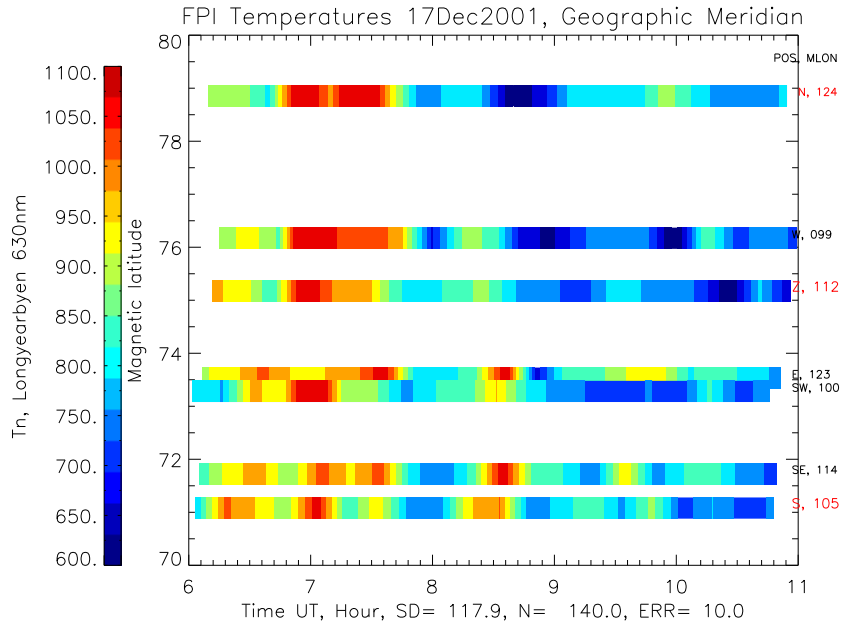


Figure 5.8: This is the FPI temperature from 17^{th} December, 2001, shown in the same format as figure 5.7.

the F-region neutral thermosphere are not a simple one-to-one relationship, which could be the case if, for example, simple cusp particle precipitation heating, in-situ, is solely responsible for the neutral temperature rise. Overall, the difference between the highest and lowest neutral temperatures within this period is as high as 500K, demonstrating a significant neutral energy contribution to the cusp energetics.

5.6.3.3 FPI LOS wind

The LOS neutral wind is shown in figure 5.5. The first comparison is between the LOS wind, neutral temperature, and red-line intensity from the zenith position only, labelled 'Z', between figures 5.5, 5.8 and 5.7 respectively. The zenith LOS wind represents direct FPI measurements of the vertical wind. Periods of upwelling of around 50 m/s can be found. Periods of the highest upwelling match well with the zenith intensity enhancements. The geographic west (W) and east (E) FPI positions show clear westward zonal flows. The zonal flows also dominate over the geographic meridian flow shown by the FPI south (S) and north (N) directions. The highest westward zonal flow is found over the highest red-line intensity period.

For a quiet FPI period, the neutral wind is expected to following the antisunward direction in general because solar pressure force dominates. The antisunward direction is approximately westward at 8 UT and becoming more northward by 14 UT as the Earth rotates. Hence, the high westward flow in neutral wind shown in case study 1 is not typical of the quiet and steady neutral condition. It is caused by the increased ion-neutral coupling from the active conditions brought upon by the southern IMF, and has possible influences from the cusp dynamics.

During the period 7-10 UT, the EISCAT radar made a 4-point 'windshield wiper' scan, measuring points 5 to 7 (figures 5.2 and 5.4) along the magnetic meridian line. Strong LOS Vi were measured along the northwest direction from 7:20-9:10 UT. The FPI measured weak winds of $\pm$60 m/s until $\sim$7:45 UT and then the winds are accelerated rapidly to nearly 300 m/s with a strong westward component. They abate and then accelerate and remain strongly westward for nearly 2 hours until 10 UT, abate for $\sim$10 minutes and accelerate. The ion velocities have dropped to 100 m/s by 9:30 UT, so the continuing large westward neutral wind is being carried by inertia. The consequences of the acceleration of the neutral gas through ion-neutral collisions is that the Joule heating is large before 8 UT when there is a large difference between ion and neutral velocities. As the neutrals

are accelerated, the Joule heating falls off and neutral temperature drops.

Using the same equations as Kosch et al. [2011] and the e-folding time described in section 2.3.4, the estimated neutral atmospheric response time ranges from 2.6 to 2.8 hours for the altitude range of 200-240km. This timing represents a general large scale result. Furthermore, as noted by Heelis et al. [2002], Kosch et al. [2011], small scale response can be as quick as 30 minutes depending on latitude, geomagnetic activity and local time. Hence, the response time observed here is possible.

5.7 Case study 2 - 18^{th} December 2001

5.7.1 MSP

The MSP data from case studies 2 are shown in figure 5.9. The cusp can again be identified within the observing period between 6 and 11 UT, with soft cusp precipitations demonstrated by regions where red to green-line intensity ratios are high. The cusp is seen more towards the northern side for case study 2, comparing to the more southern cusp in case study 1.

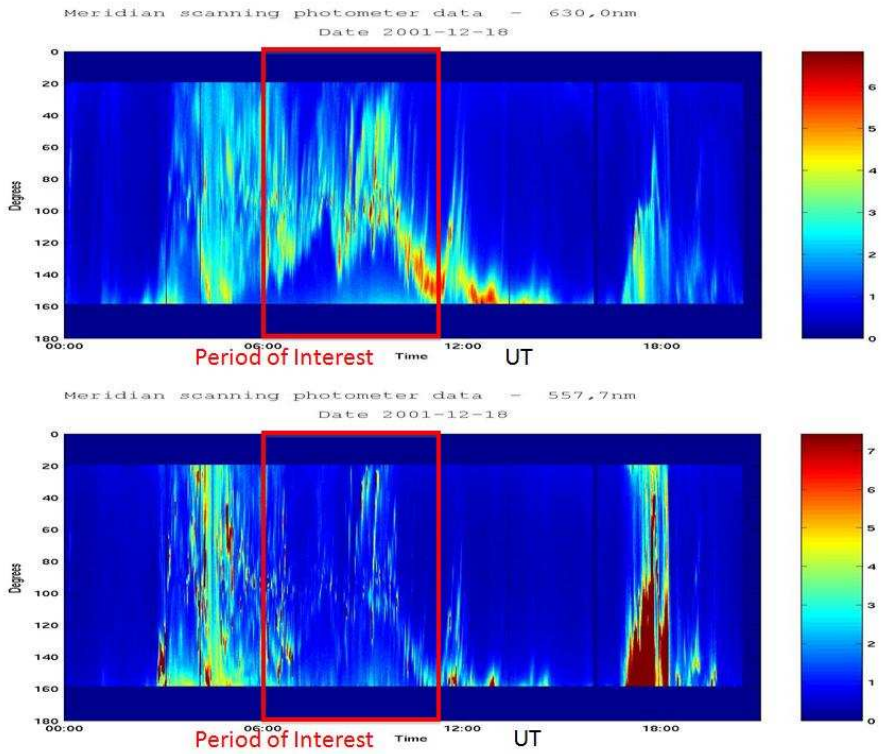


Figure 5.9: The top and bottom plots are the red (630.0 nm) and green-line (557.7 nm) Ny-Alesund MSP intensity data for 18^{th} December 2001.

5.7.2 EISCAT

The EISCAT data from case studies 2 are shown in figure 5.10. The F-region EISCAT Ne, Te, Ti and LOS Vi are again displayed in the figure. The increased F-region Ne also matches well with the 630 nm cusp intensity variations shown by the MSP. Electron density for case study 2 is lower than the values in case study 1, at smaller or equal to $10^{11.5}$ m^3. The highest temperatures for case study 2 are similar to case study 1 and reach 1500K for the cusp regions with increased electron density, but are found more on the northern side. A northwestward flow is also seen at 8 and 9 UT for case study 2, but for a shorter time period than case study 1.

5.7.3 FPI

5.7.3.1 FPI intensity

As explained in Case study 1, the FPI and MSP observe a similar range of latitudes so figures 5.13 and 5.9 may be compared directly. The FPI and MSP intensities both show a clear red-line intensity enhancement caused by cusp precipitation. It can also be seen that the highest cusp red-line intensity is found more towards the northern side for case study 2, when compared to case study 1.

5.7.3.2 FPI temperature

The FPI temperatures in figure 5.14 can again be compared to the FPI intensities in figure 5.13. The most distinctive feature is that the regions of highest neutral temperature enhancements do not overlap with the regions of highest intensity enhancements. This feature is also clearly demonstrated by case study 1. The highest temperature and intensity regions are divided between north and south in case study 2, comparing to the zonal/time divide in case study 1. Case study 2 re-confirms that the temperature and intensity enhancements in the F-region neutral thermosphere are not a simple one-to-one relationship.

5.7.3.3 FPI LOS wind

The LOS neutral wind is shown in figure 5.11. Periods of upwelling of around 50 m/s can be again be found, similar to case study 1. However, the upwellings in case study 2 are more prolonged and less clearly linked with zenith intensity enhancements. The W and E

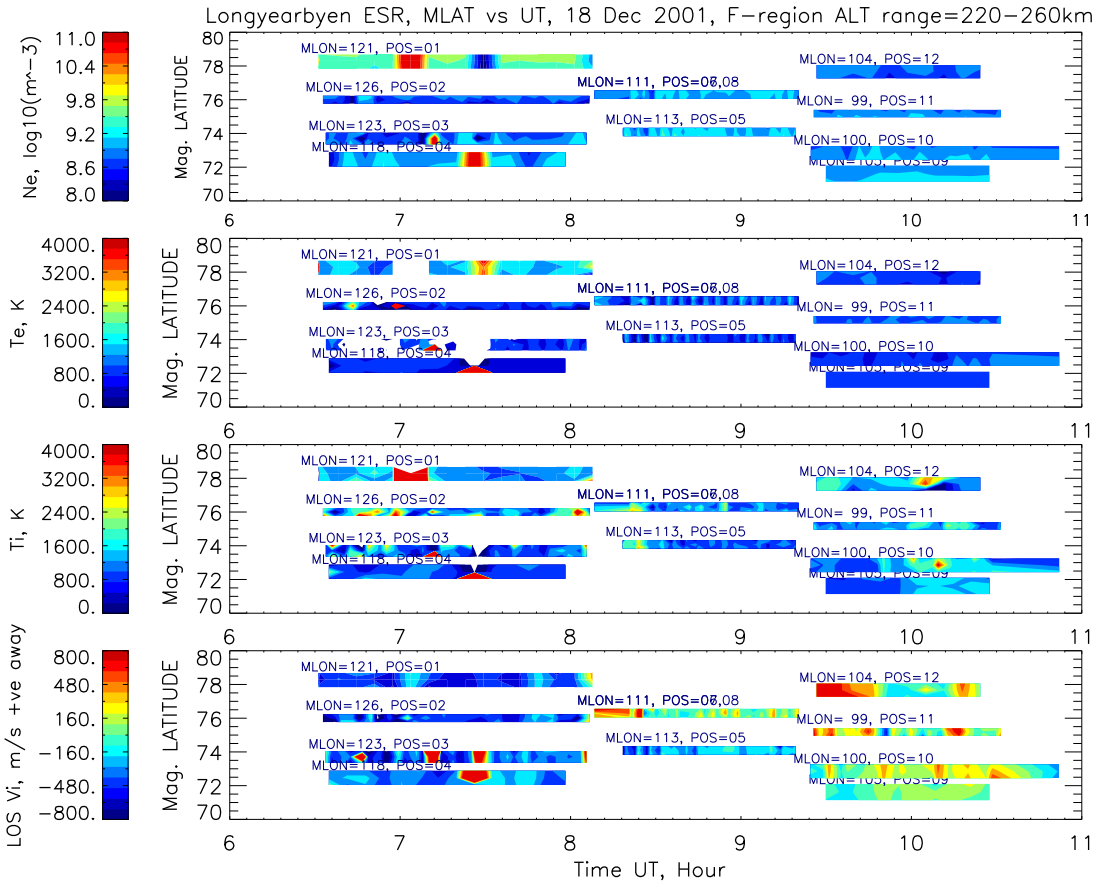


Figure 5.10: This is similar to figure 5.4, but data from 18^{th} December 2001 are shown. Note that although the radar settings for both nights are similar, the positions and timings are slightly different, especially for the meridional scanning period between 8-9 UT for this night.

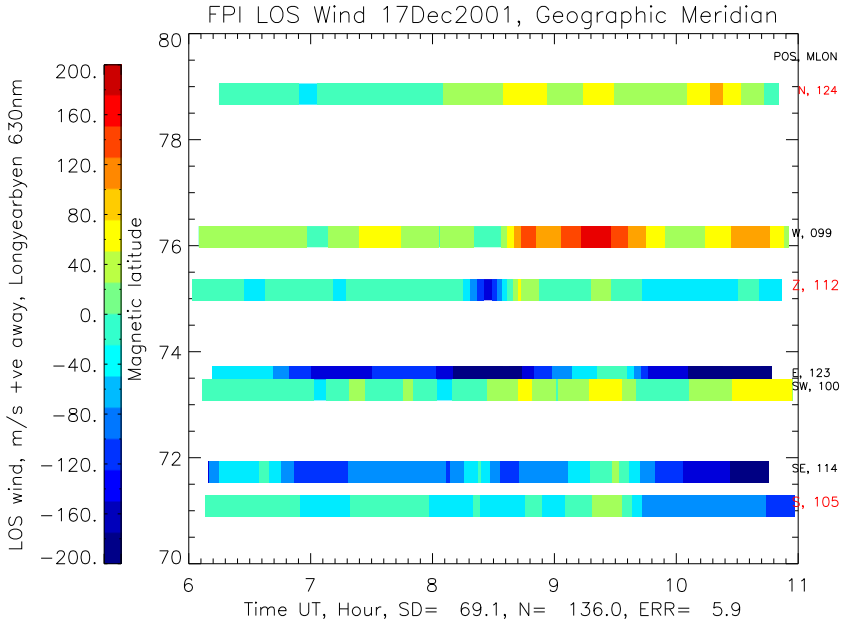


Figure 5.11: This is the FPI LOS neutral wind from 18^{th} December, 2001, shown in the same format as figure 5.7.

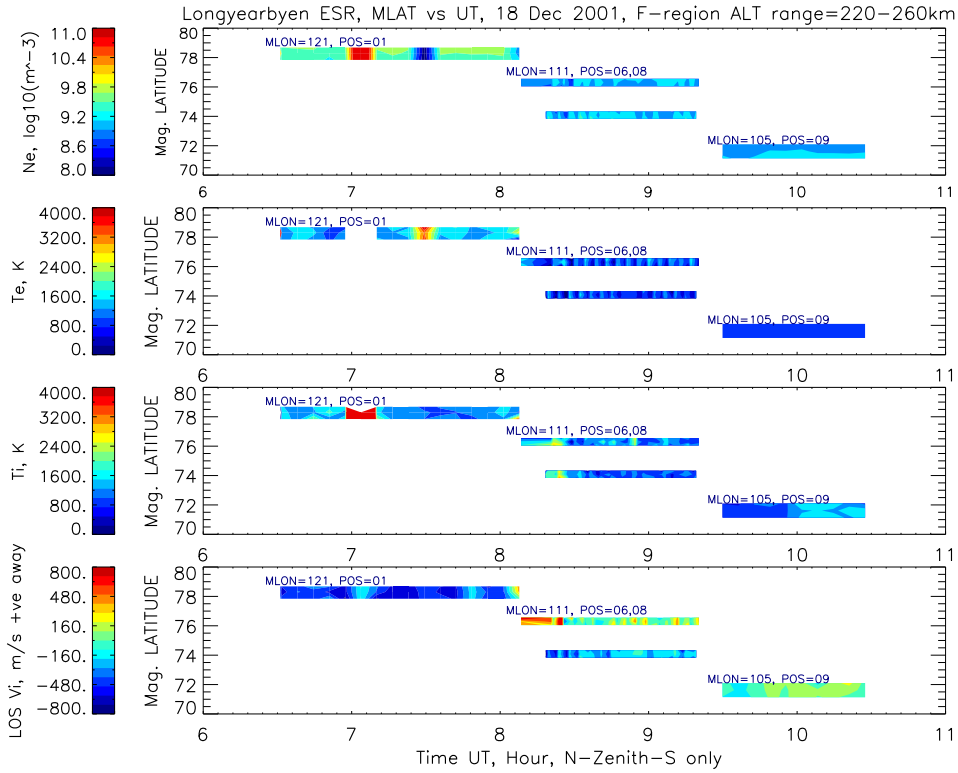


Figure 5.12: This figure is similar to figure 5.4, but only the geographic north-zenith-south directions are shown.

FPI positions also show clear westward zonal flows. A small northward flow can be found in case study 2 after 9:30 UT. It gives a combined approximate direction of north-westward. The high westward flow in neutral wind shown in both case studies is not typical of the quiet and steady neutral condition and has possibly influences from the cusp. Strong LOS Vi were measured along the northwest direction by EISCAT (figure 5.10) between 8 and 9 UT. A similar argument as case study 1 can therefore be made about the spatial distribution of neutral temperature and red-line intensity, with ion-neutral collisions and the difference in ion velocity and neutral wind (section 5.6.3.3).

5.7.4 Case studies 1 and 2 discussion and summary

The first two case studies demonstrate that the F-region neutral temperature can vary and reach 500K higher than background under the active and intense IMF southward cusp conditions. The cusp location has been confirmed by the MSP/keogram data and it is found within the FPI's FOV. The ion velocities show a general strong geographic northwest (geomagnetic north) flow between 7 and 10 UT for case study 1, and between

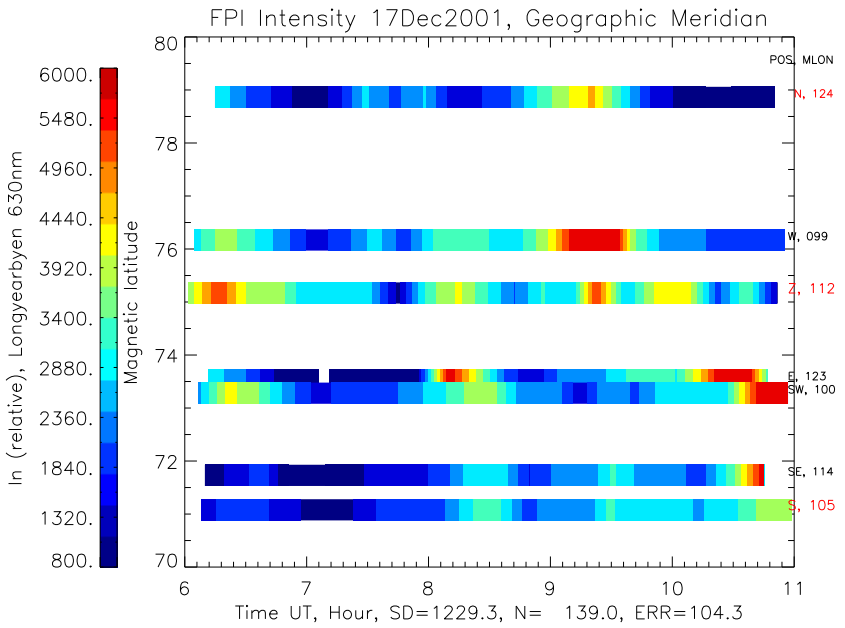


Figure 5.13: This is similar to figure 5.7, but 630.0nm intensity from 18^{th} December 2001 is shown.

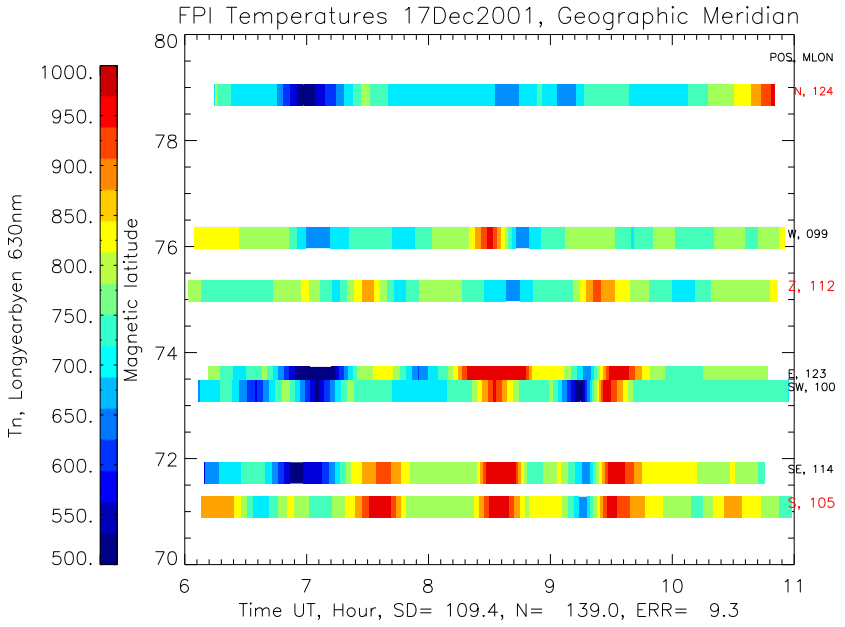


Figure 5.14: This is the FPI temperature from 18^{th} December 2001, shown in the same format as figure 5.7.

8 and 9 UT for case study 2. The neutral wind is affected by the ionospheric conditions in both cases and show a westward component in the LOS wind, instead of the expected northward flow only, for quiet and steady conditions. The highest neutral temperature enhancement region is found adjacent to the highest red-line intensity region and do not directly overlap. Overall, the high neutral temperature enhancement is expected to have significant effect on the density rise seen by satellite data above the cusp regions.

5.8 Case study 3 - 10^{th} January 2008

The geomagnetic condition for case study 3 was quiet with Kp smaller or equal to 3 over the experiment period, in contrast to the active data for case studies 1 and 2. The data are taken from 10^{th} January 2008 under solar minimum and IMF By dominant conditions, also in contrast with the southward Bz IMF conditions for the other case studies. All three cases are therefore valuable for the interpretation of the statistical FPI data, which are separated into samples under solar maximum/activity and minimum/quiet average conditions. It should be noted that the plots for case study 3 are shown in MLT, while the other cases are described in UT. MLT is used because it is more appropriate for cusp dynamics, which are closely associated with geomagnetic and magnetospheric variations. MLT=UT+3.1 hours for the Longyearbyen site.

SCANDI data, instead of FPI data, are used for this case study to provide the neutral measurements. The FPI uses an observing mirror that cycles through the look directions, but SCANDI applies a fisheye lens and it is able to measure all-sky maps of the neutral parameters. The SCANDI all-sky map, as well as the MSP and EISCAT look directions, projected at 240km altitude, are shown in figure 5.15. The SCANDI FOV was divided into a 25-zone map at the time of the experiment, and each zone is able to provide an individual measurement of the neutral intensity, temperature and LOS wind. A horizontal wind field can also be fitted because of the map measurements.

EISCAT was measuring the geographic north and south directions at 60 degrees zenith from the 32m radar dish, and the field-aligned direction from the 42m dish. The north and south EISCAT directions overlap with the north and south SCANDI zones in the outer ring (zones 15 and 21), while the field-aligned direction overlaps with SCANDI zenith zone. The setup therefore restricts the observation to the geographic meridian for both instruments. There is also a spatial difference in between SCANDI/EISCAT and the MSP measurements, which sweeps along the geomagnetic meridian.

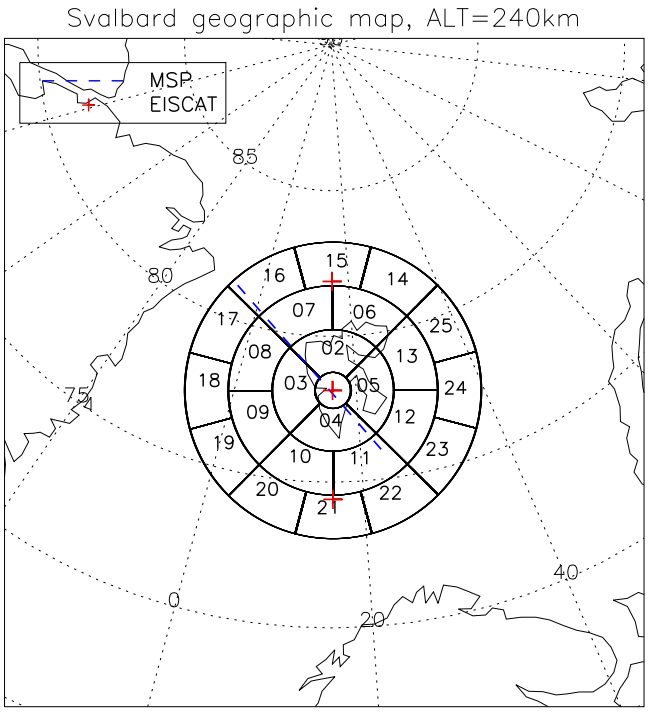


Figure 5.15: This map shows the relative positions between the SCANDI 'zones', the three EISCAT observing points (red crosses) and the MSP 'windshield scanning' position projected at 240km altitude (blue dash line). This map is appropriate for case study 3 on 10^{th} January, 2008.

5.8.1 MSP

The red and green line emissions from the MSP are shown in figure 5.16. The MSP data are restricted to the time range when both SCANDI and EISCAT were in operation. The cusp can be clearly identified on the northern side of its FOV over this time period. The northward cusp location is as expected for the quiet and small IMF Bz and By dominant conditions. The highest soft cusp precipitation is found at around 11 MLT, where the red-line to green-line intensity ratio is also the highest. Three magnetic latitude axes are given on the right. These indicate the location of the emission of the red-line emission peak is at an low altitude of approximately 200-210 km (section 5.8.3) around 77 degrees magnetic latitude.

5.8.2 EISCAT

The EISCAT parameters are shown in figure 5.17. Unlike the EISCAT plots for case studies 1 and 2 (figures 5.4 and 5.10), which only show EISCAT data from the F-region altitude ranges; the figure for case study 3 includes the whole altitude range observed by

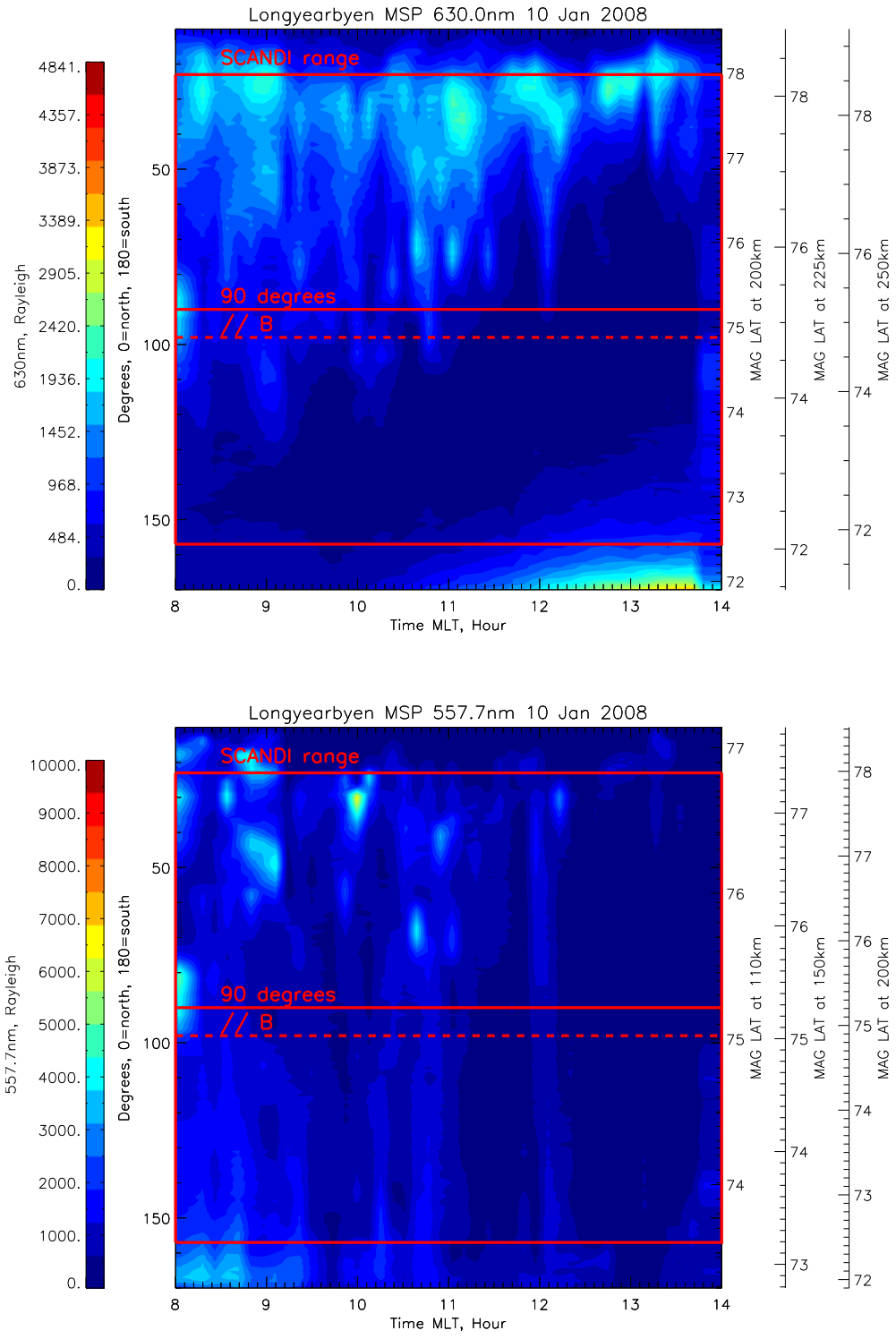


Figure 5.16: The top and bottom plots show the red-line 630.0nm and green-line 557.7nm MSP emission intensities respectively. The right hand side includes two addition axes, which demonstrates the appropriate magnetic latitude range if the peak emission intensity occurs at altitudes of 200, 225 and 250 km for red-line, and 110, 150 and 200km for green-line. The SCANDI range, 90 degrees (zenith) line, and parallel to magnetic field lines are shown in red.

each of the radar's look directions. The figure is plotted in magnetic latitude vs. time, which helps to track the movement of the cusp over a wider latitude range.

The electron density data (top panel in figure 5.17) are compared with the red-line MSP data (top panel in figure 5.16). The high electron density spot of $10^{11.0}$ m^{-3} at 10 MLT is associated with the harder precipitations that affect both F and E-regions. Softer cusp precipitations at 11 MLT can also be identified by electron density rise to $10^{10.8}$ m^{-3} and the red-line intensity enhancement from the MSP data. The electron densities are higher on the northern EISCAT FOV, which also matches spatially with the cusp activity towards the north.

High electron and ion temperatures at 3500K are also detected at this soft cusp precipitation period at 11 MLT. It should be noted that the continuous band of greater than 4500K electron temperature near the centre FOV comes from the high electron temperature at high altitudes, detected by the field-aligned radar direction, and are not cusp related; similarly for the continuous band of LOS ion velocity at higher than 400 m/s. Both are likely to be noise. Spots of high northward ion velocity are detected at around 10 to 11 MLT, at the time period when cusp precipitations and electron densities are high.

5.8.3 The peak red-line emission altitude

The electron density pattern (top panel, figure 5.17) can be compared to the MSP red-line intensity pattern and the three magnetic latitude axes appropriate for altitudes of 200, 225 and 250 km in the MSP plot (figure 5.16). The MSP intensity and EISCAT electron density can be roughly approximated to be proportional to each other. The elevated electron density centres at 11 MLT from EISCAT is associated with $\sim$77 degrees latitude. This feature can be linked to the MSP red-line intensity increase also at 11 MLT. It can be seen that the magnetic latitude axis for 200km altitude provides the best match of $\sim$77 degrees for the increased intensity region. The comparison indicates the peak red-line emission is also at approximately 200-210 km altitude, under extremely quiet conditions.

5.8.4 SCANDI

5.8.4.1 SCANDI intensity

The SCANDI red-line intensity data are shown in figure 5.18, and it can be compared to the MSP red-line intensity in figure 5.16. The intensity patterns match well; with consideration that geographic meridian SCANDI data are used, while the MSP observes

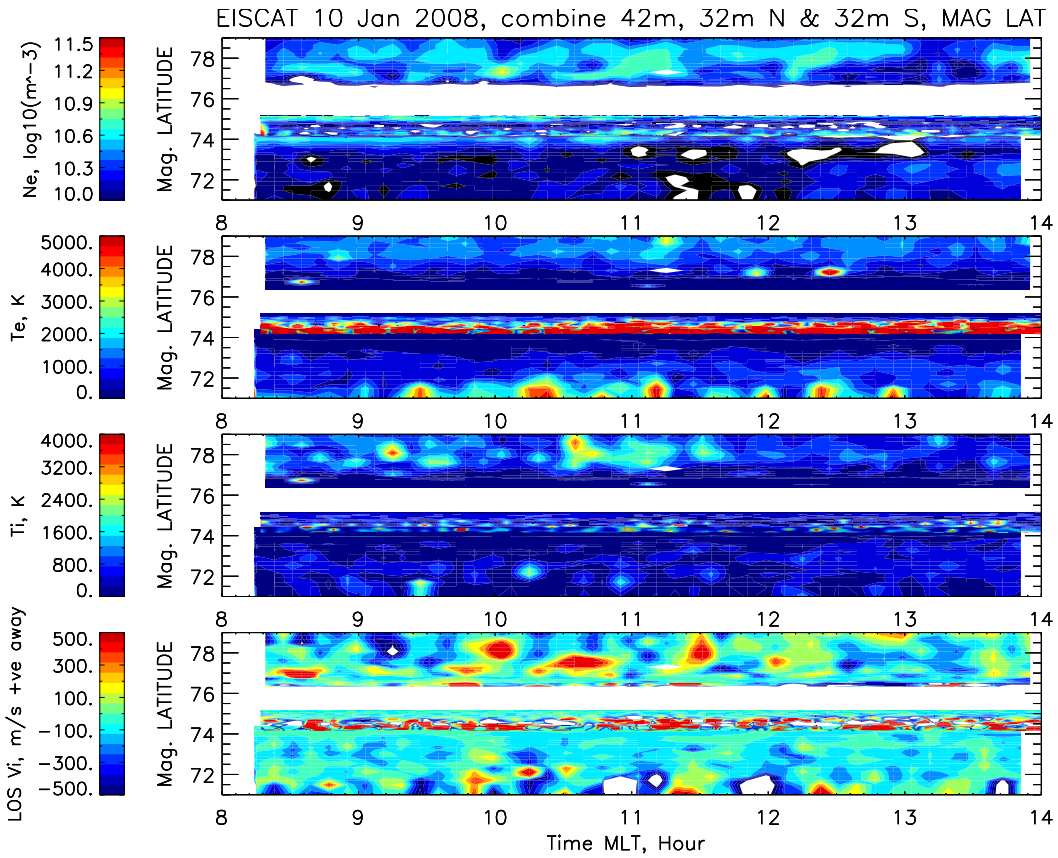


Figure 5.17: These are the EISCAT data from all three observing positions as shown in figure 5.15. All four common EISCAT ionospheric parameters are displayed: Ne, Te, Ti and LOS Vi from the top to bottom panels. The y-axis is plotted in magnetic latitude, where the whole altitude ranges for all three radar pointing directions are shown. This is different from e.g. figure 5.10, where only F-region altitudes are shown. The top 'band' therefore contains data from the northward-pointing position of the 32m dish, and the bottom 'band' from the southward-pointing position, also from the 32m dish. The narrow stripe/band close to the middle are data from the 42m dish, which points along the magnetic field and covers the smallest latitude range.

along the geomagnetic meridian and has higher time and spatial resolutions. The red-line intensity is the most enhanced for the soft cusp precipitation period at 11 MLT.

5.8.4.2 SCANDI temperature

The neutral temperature data are shown in figure 5.19. Comparing the neutral temperature to the SCANDI red-line intensities (figure 5.18), the intensities are the most increased by cusp precipitations towards the north and between 8-12 MLT. The highest temperature enhancement regions do not overlap, and are found towards the south and at closer to zenith after 12 MLT. The non-overlapping intensity and temperature regions are similarly

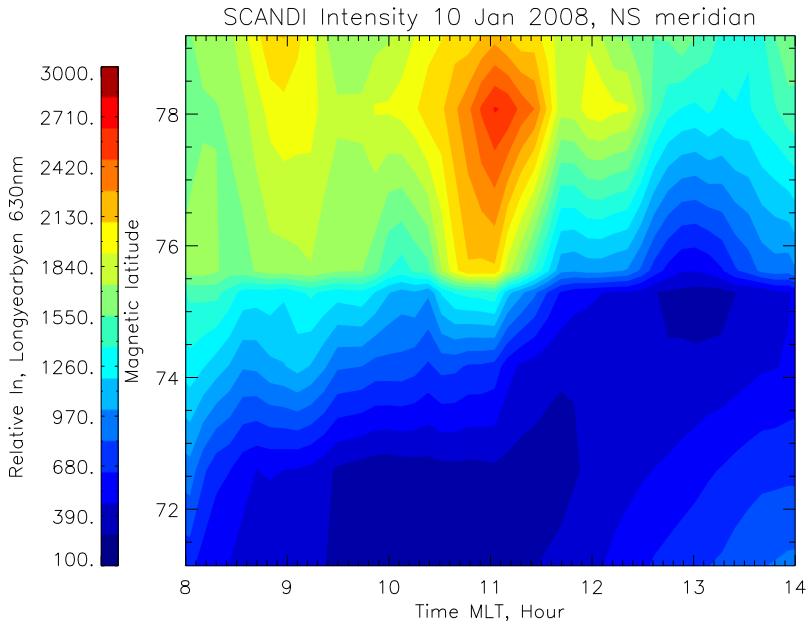


Figure 5.18: This plot shows the SCANDI relative red-line intensity data, taken from zones along the geographic meridian (figure 5.15). The data along the magnetic latitude axis are constructed from measurements at zone 15, average of zones 6 and 7, zone 2, zone 1, average of zones 10 and 11, and zone 21. The geographic meridional data are shown in a 'keogram' format.

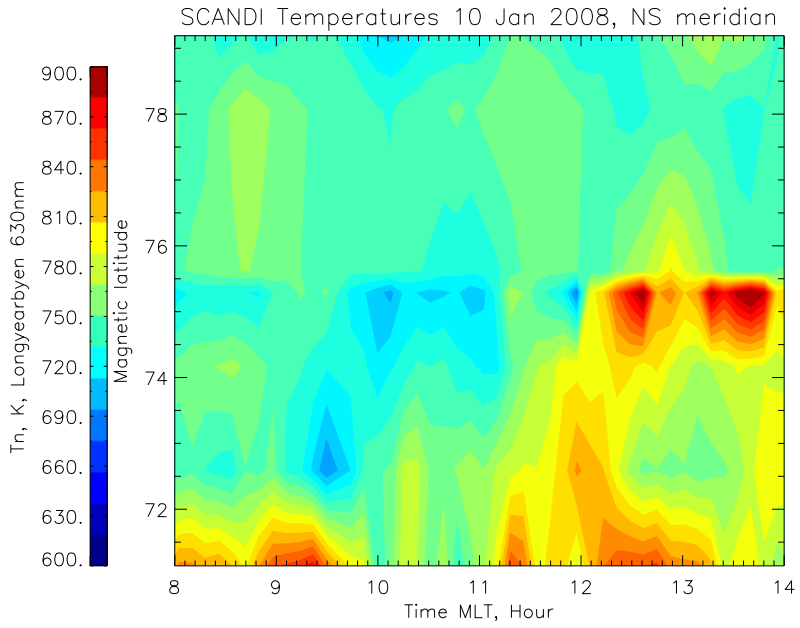


Figure 5.19: This plot shows the SCANDI F-region neutral temperature along the meridian zones and is similar to figure 5.18.

found in case studies 1 and 2. The temperature is raised by about 200K in this solar minimum and quiet case study, while the more active and solar maximum cases 1 and 2 show temperature rises of as much as 500K. The radar temperature data do not show similar temperature increase in the case study here, and checks with magnetometer data also indicate there are no matching currents that may suggest the neutral temperature increase is caused by other geomagnetic influences.

The higher spatial resolution SCANDI data here are also able to show an additional temperature rise of 50 K at latitudes above 75 degrees and between 11 and 12 MLT. This appears to spatially match with the highest intensity enhancements at the same latitudes about half an hour earlier. The smaller 50K temperature rise here is also accompanied by electron and ion temperature enhancements in the EISCAT data (second and third panels in figure 5.17).

To summarise, the highest temperature enhancement is found at regions next to/ around the highest cusp precipitation region. This result is similar to those found in case studies 1 and 2. Furthermore, the higher spatial resolution SCANDI data are able to show a smaller temperature rise in a region matching with the precipitation region, with an approximate half hour time delay.

5.8.4.3 SCANDI wind

The zenith wind data are shown in figure 5.20, and upwelling reaching 40 m/s is detected in the highest temperature rise region, demonstrating neutral gas upwelling associated with heating. SCANDI's capability of measuring instantaneous maps of neutral parameters also enables the LOS winds to be fitted into a horizontal wind field, and this is shown in figure 5.21. The neutral wind under quiet and steady conditions is expected to follow the anti-sunward direction, because of the dominating solar-pressure force. It can be seen that the neutral wind in general follows the anti-sunward blue line in the figure very well. The neutral wind then deviates slightly from the anti-sunward direction between 9:15 and 10:15 MLT, where the neutral wind has a more westward component than the anti-sunward direction. In comparison, stronger and more prolonged westward neutral winds are shown by case studies 1 and 2. The smaller cusp influence on the neutral wind is appropriate for the much quieter case study here. The ionospheric convection pattern is unlikely to be the simple 'twin-cell' pattern during this strong By and small Bz period. Overall, the neutral wind flow is mostly unaffected by the cusp dynamic over this quiet cusp period.

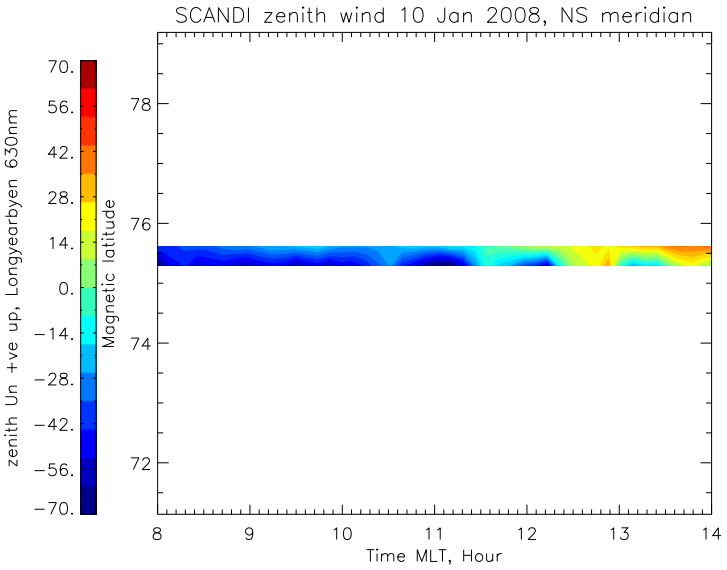


Figure 5.20: This plot shows the SCANDI LOS wind from the zenith zone and the two meridian zones from the inner 'ring' i.e. zones 1, 2 and 4 in figure 5.15. The vertical neutral wind component therefore dominates.

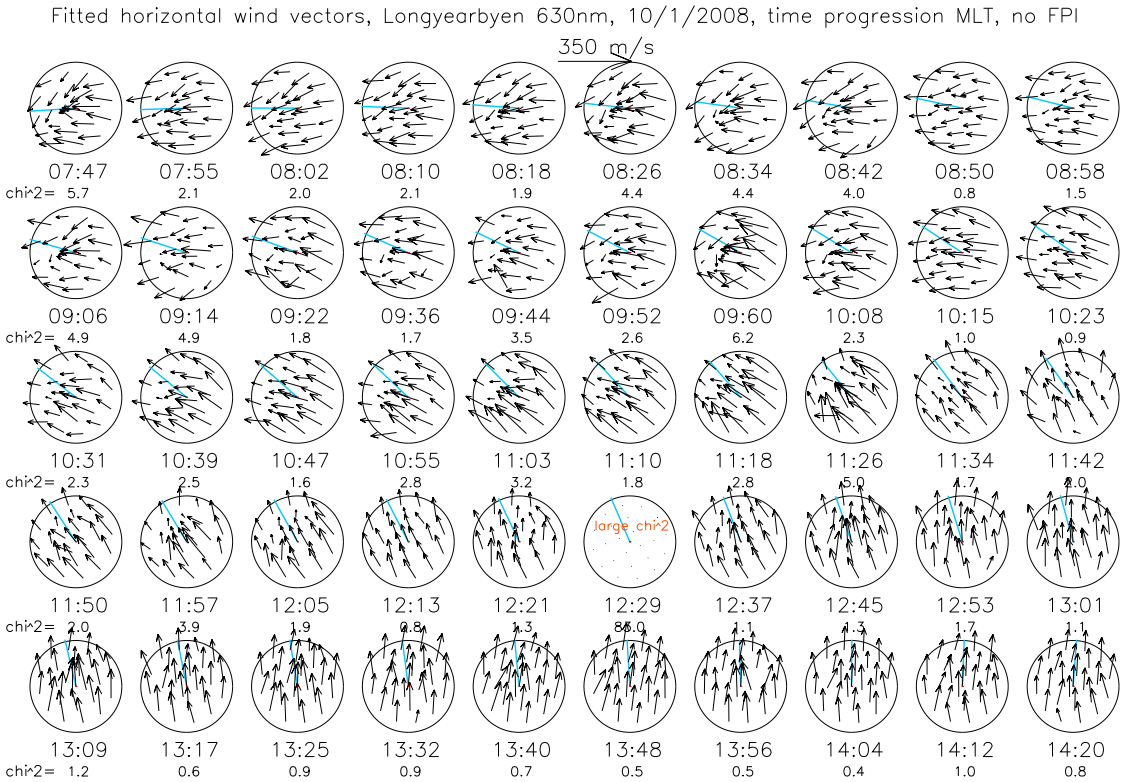


Figure 5.21: This figure shows the fitted SCANDI horizontal wind field, where the blue line in each plot points towards the antisunward direction. Top of each plot is geographic north, right is east and so on. The χ^2 value for each plot is also shown, which monitors the quality of fit for each vector field. Bad quality fits with large χ^2 are excluded.

5.8.5 Case study 3 discussion and summary

Case study 3 demonstrates results that are complementary to the first two case studies, where all three cases indicate the highest temperature rise regions are found next to/ around the high cusp precipitation regions. The higher spatial resolution for this case also show an additional but smaller temperature rise of about 50 K in a region close to the highest cusp precipitation region. The neutral wind flow in general is unaffected by the cusp and maintains an anti-sunward flow, apart from a small westward component for a relatively short period of time. This feature is also seen in the higher activity period for the first two case studies, where a much stronger westward flow is seen when a northerly antisunward flow is expected. The cusp here is detected at higher magnetic latitude than the other two case studies, as expected for the quieter and small Bz condition here.

5.9 Application and purpose of the case studies

The three case studies are in support of the statistical FPI results that follow this section. They are able to confirm the cusp is often within the FPI and SCANDI FOV from the MSP and EISCAT data. Furthermore, the neutral atmosphere is affected and especially the temperature is increased by the cusp precipitations. Upwelling is also detected, as well as a westward zonal flow. The case studies are also able to confirm the relative locations of the rise in neutral temperature and the high cusp precipitation regions in the individual cases, which are determined under different solar, IMF and geomagnetic conditions. The detailed ionospheric dynamics are beyond the scope of this study, and they can be found in investigations that focus on magnetosphere-ionosphere interactions, such as Carlson et al. [2006], Moen et al. [2004], Walker et al. [1999], Rodger et al. [1995], Lockwood and Carlson [1992].

Both EISCAT experiments from case studies 1 and 2 were specifically designed to observe the cusp. Case study 1 has already been used for an ionospheric patch study by Carlson et al. [2006], where the cusp was identified. Furthermore the keogram and MSP data for all three case studies show clear cusp features. Hence, it is unlikely that the cusp was not in the location as described in this thesis.

5.10 Statistical results from solar maximum and minimum years

The average statistical FPI data are shown in this section, where the three case studies provide additional information to confirm and explain the average neutral features seen under cusp influences. The aim of the statistical data study is to deduce the most common neutral reactions to cusp precipitations under different solar and geomagnetic conditions. The analysis could help to identify additional neutral contributions to the already-existing cusp studies that focus on the magnetosphere and ionosphere only e.g. Moen et al. [2004], Walker et al. [1999], Rodger et al. [1995]. It can also add to the reasons for densities enhancements seen in the high altitude regions e.g. Clemmons et al. [2008], Demars and Schunk [2007], Schlegel et al. [2005], Lühr et al. [2004].

The Svalbard FPI has a 1 degree FOV and pointing direction at an elevation angle of 30 degrees. The north and south look directions are separated by ∼830km. A peak emission height of 240 km for the red-line centred at geomagnetic location 78N, 156E would have a FOV of ∼10 degrees latitude.

Please note the different panels shown in the figures in this section each have their own colour scale, appropriate for the average background geomagnetic conditions they each represents. For example, two panels under quiet and moderate geomagnetic conditions will have background magnitudes quite different from each other, and using the same colour scale will not allow the features to be identified. All colour scales are clearly labelled.

5.10.1 Red-line intensity, In

5.10.1.1 Solar maximum years (1999-2004)

The average solar maximum FPI red-line intensity results are shown in figure 5.22, top row. From the intensity data, the red-line intensity enhancements caused by the cusp are located on the southern side of the FPI FOV. The cusp remains at the south for both quiet (Kp=0-3) and moderate (Kp=4-7) geomagnetic conditions. Active geomagnetic condition data (Kp greater than 7) are not shown because the sample size can be smaller than 2 nights. The intensity for the moderate case is higher, representing higher cusp precipitation density, as would be expected for greater activity conditions.

5.10.1.2 Solar minimum year (2005-2008)

The average solar minimum FPI red-line intensity results are shown in figure 5.23, top row. The cusp is located on the northern side of the FPI FOV, in contrast to the solar maximum results. The cusp also remains at the north for both quiet and moderate geomagnetic conditions. However, the intensity in the middle strip, representing zenith position, indicates an extension of the cusp towards the south for moderate activity conditions. Note that the absolute intensity value in figure 5.23 (solar minimum data) is not comparable to figure 5.22 (solar maximum data) because of instrumentation hardware changes.

5.10.1.3 Discussion

In general, previous studies show that the red-line 630.0 nm emission dominates the dayside aurora, and the optical detection of the cusp is about 5-8 hours wide, centred around magnetic noon, and with a 100-200km latitudinal width [Smith and Lockwood, 1996, Lockwood and Carlson, 1992, Sandholt et al., 1986, Sivjee et al., 1982, Sandholt et al., 1980, Deehr et al., 1980]. The position of the cusp depends on the shape and position of the magnetopause. It is controlled by a combination of both the nightside current system, and has its own weaker but separate current system on the dayside, controlled by IMF Bz. Hence it is expected that there are equatorward movements of the cusp when IMF Bz turns southward. As IMF Bz turns north, the opposite happens and the dayside auroral oval also moves poleward. This IMF dependence is well documented in e.g. Moen et al. [2004]. It can be assumed that the cusp follows the latitudinal position of the auroral oval.

The average southward/northern position of the cusp intensity detected by the FPI is therefore as expected for the solar maximum/minimum conditions from earlier studies. Case studies 1 to 3 also give support that the cusp is detected towards the south/north of the FPI FOV at solar maximum/minimum conditions. The movement of the cusp under different solar and geomagnetic conditions over the FPI FOV means the FPI provides a window on the averaged thermospheric conditions, on the centre, as well as the northern and southern boundaries of the cusp.

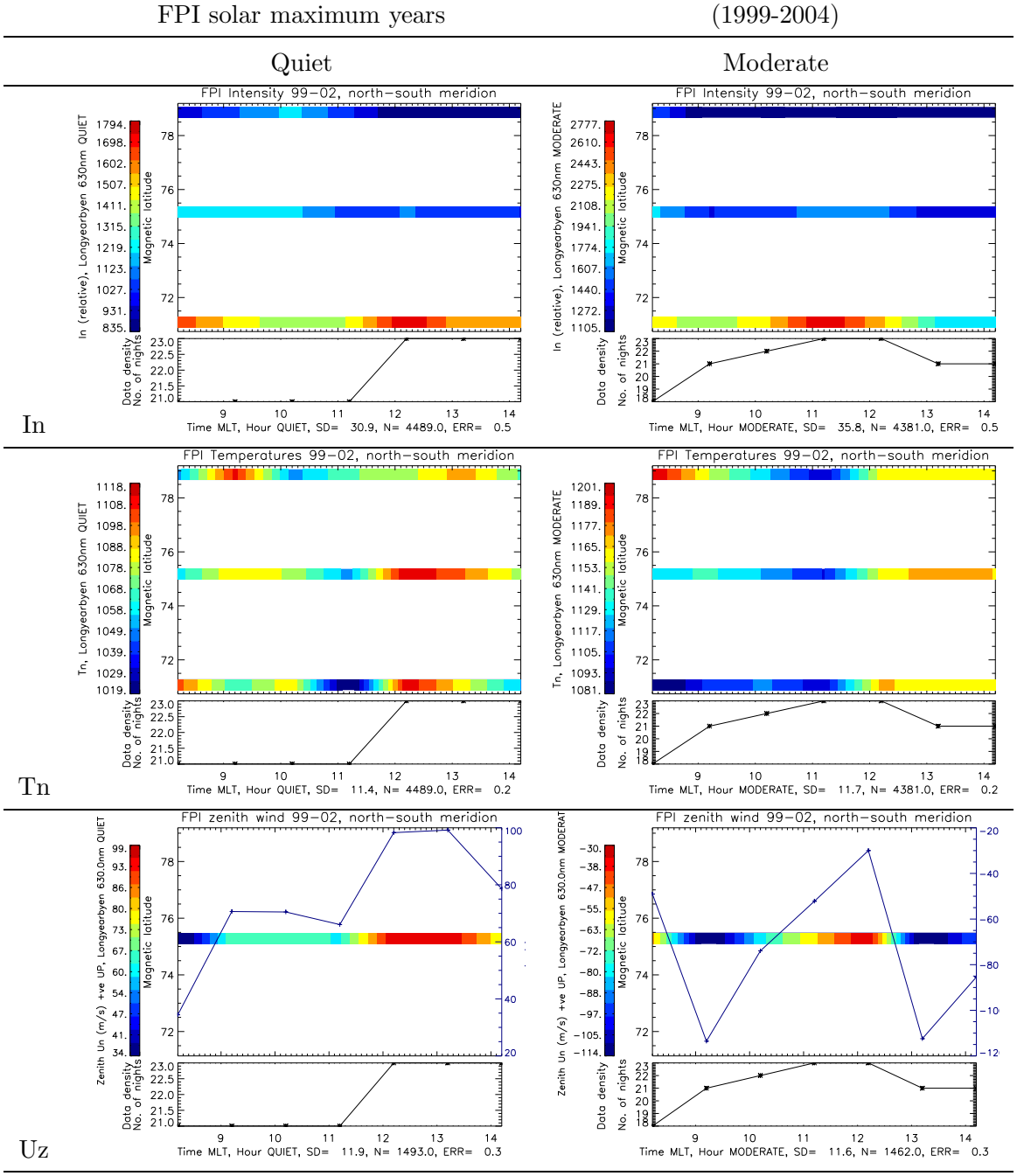


Figure 5.22: These are the average FPI data from solar maximum years of 1999-2004. Data from 8-14 MLT are shown. The left column contains data from quiet geomagnetic conditions with Kp of 0 to 3, and the right column contains data from moderate geomagnetic conditions with Kp 4-7. The rows from top to bottom are the red line intensity (In), neutral temperature (Tn) and zenith wind (Uz). Each plot includes a data density graph showing the number of nights of data that are used for each hour in UT. The standard deviation (SD), number of data points (N) and error (ERR) calculated with SD/sqrt(N) are also noted.

5.10.2 Neutral temperature Tn

5.10.2.1 Solar maximum years (1999-2004)

The average solar maximum FPI temperature results are shown in figure 5.22, middle row. Overall, the temperature is enhanced to above 1000K over the whole FOV and time period, compared with the usual background value of 700-800K. Comparing the quiet and moderate activity temperature plots, the quiet activity temperatures are more spread out in both latitude and time, and reach values of above 1100K; while the moderate activity temperatures are also enhanced over the whole FOV. The highest values of temperature are found in regions adjacent to the highest intensity area, reaching 1200K for moderate activity results, at about 100K higher than the quiet activity value at solar maximum.

5.10.2.2 Solar minimum years (2005-2008)

The average solar minimum FPI temperature results are shown in figure 5.23, middle row. Overall, the temperatures are at lower values between 800 and 1000K. For the quiet activity results, the highest temperature regions at 1000K are detected to the south; again the high temperature clearly do not overlap with the highest intensity region at the north under solar minimum conditions. For the solar minimum higher activity results, the temperature enhancements are more spread out and overlap with the intensity enhancements, and reach maximum values of nearly 1100K.

5.10.2.3 Discussion

Comparing the average temperature results to those from the case studies, the feature of highest temperature regions not overlapping with the highest cusp intensity region can be seen in all three case studies. Smaller temperature rises which overlap with the high intensity spots are also seen in the quiet activity case study 3 and the statistical solar minimum-moderate activity average, as well as the solar maximum-quiet activity average.

Overall, the statistical results here can be seen as 'windows' at different latitudinal regions around the cusp, with the reasoning that the latitudinal location of the cusp approximately follows the location of the auroral oval. Its location is therefore dependent on IMF Bz which is directly linked to geomagnetic activity levels, as well as the overall solar activity. Therefore, the FPI FOV appears to cover the most northern boundary of the cusp as shown in the solar maximum-moderate activity plot. The solar maximum-

quiet activity plot covers more of the northern boundary. Then the FOV covers the southern side of the cusp in the solar minimum-moderate activity plot; and finally, the solar minimum-quiet activity plot sees the most southern boundary of the cusp.

Comparing the relative intensity and temperature regions, when the FPI is observing the extreme northern and southern sides (i.e. solar maximum-moderate activity and solar minimum-quiet activity) of the cusp, the most obvious feature is the lack of overlaps of intensity and temperature regions. In contrast, when the FPI is observing closer to the cusp centre (i.e. solar maximum-quiet activity and solar minimum-moderate activity), some overlap of the temperature and intensity regions can be seen, although the raised temperature region remains more spread out then the raised intensity region. The interpretation of these plots is complicated because it is not as simple as moving a telescope to view different sections of the cusp. For example, the solar maximum cusp conditions are different from the solar minimum, because there may be statistically different trends in the underlying geomagnetic and solar conditions involved.

5.10.3 Vertical Un

5.10.3.1 Solar maximum years (1999-2003)

The vertical wind data at solar maximum are shown in figure 5.22, bottom row. Upwelling can be seen over the whole time period in the solar maximum-quiet activity case and reaches high values of 100 m/s at 12-14 MLT. The region of highest upwelling also matches with the region of highest temperature. However, downwelling is also seen in the solar maximum-moderate activity case, reaching values of -100 m/s.

5.10.3.2 Solar minimum years (2005-2008)

The vertical wind data at solar minimum are shown in figure 5.23, bottom row. Small upwelling of 10 m/s is detected at 11 MLT in the solar minimum-quiet activity case, while small downwelling of -20 m/s is also seen throughout the majority of the observing period. For the solar minimum-moderate activity case, large upwelling reaching 100 m/s is detected at 12-14 MLT, while a relatively small downwelling of -20 m/s is also seen at around 10-11 MLT.

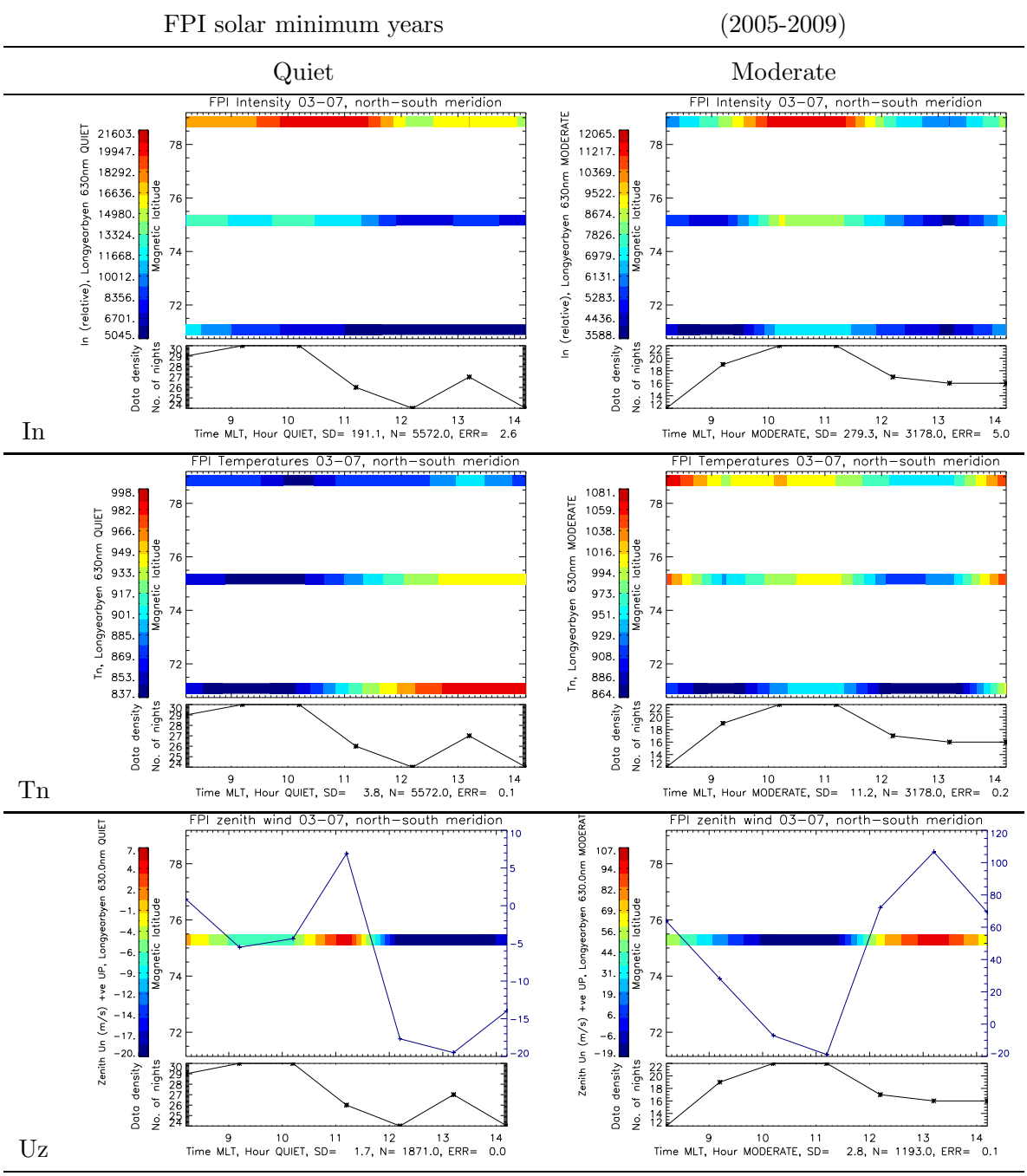


Figure 5.23: The format of this plot is the same as figure 5.22, but average data from the solar minimum years of 2005-2008 are shown.

5.10.3.3 Discussion

The vertical wind results can also be discussed with a similar argument as the temperature results. The vertical wind data can be compared with the intensity plots, where the intensity plots indicate the location of the optical cusp centre. When the FPI FOV is observing the most northern and southern sides of the cusp i.e. the solar maximum-moderate activity and solar minimum-quiet activity plots respectively; high downwelling of -100 m/s is seen in the most northern side of the cusp, while a smaller downwelling of -20 m/s is also detected for the majority of the time period for the most southern boundary of the cusp.

When the FPI FOV is observing regions closer to the cusp centre i.e. the solar maximum-quiet activity and solar minimum-moderate activity plots; upwelling reaching 100 m/s dominates the observing period for the solar maximum-quiet case, and a smaller upwelling of 20 m/s is also seen for the majority of the time period for the solar minimum-moderate case.

The FPI data therefore confirm that upwelling in the F-region neutral thermosphere is found in the cusp region, while downwelling can occur in areas towards the northern and southern boundaries of cusp. The centre of the cusp is dominated by upwelling. The upwelling/downwelling is likely to be responsible for the increased/decreased cusp density structures seen by the CHAMP satellite [Lühr et al., 2004], as denser air from below rises and descends.

5.10.4 High spatial resolution neutral data

Finally, the high spatial resolution SCANDI data are shown in figure 5.24. SCANDI data are only available from years 2007-2009, which covers the recent extended solar minimum of cycle 23/24. Only quiet activity data are available due to the extremely low activity level. The regions covered by SCANDI and the FPI are the same, but SCANDI is able to make all-sky maps of neutral parameters at horizontal resolutions of a few hundred of kilometres.

The higher spatial resolution data confirm the detection of the cusp at the northern side of the instrument from the intensity data. This indicates that the SCANDI FOV is observing the southern boundary of the contracted cusp. The temperature data demonstrate a smaller temperature enhancement, which overlaps with the intensity enhancement region between 74 and 79 degree magnetic latitudes. The highest temperature rise is found

at the southern side of the increased intensity region between 71 and 74 degrees magnetic latitude, at the boundary of the optical cusp. The temperature range is only between 700-800 K for this extremely quiet time period.

Upwelling reaching 28 m/s is also seen, while the majority of the time period is covered by downwelling, reaching value of -40 m/s. The quiet SCANDI data can also be interpreted as an observation of the southern boundaries of the cusp, when the cusp location has contracted to the north of the instrument's FOV. The high-resolution data here therefore match well and support the results from the solar minimum-quiet FPI averages and the the quiet case study. SCANDI spatial resolution is more suited for the spatial scale of cusp dynamics and it is able to provide a more detailed description of the cusp region. SCANDI data were used successfully for a joint EISCAT-FPI-MSP cusp experiment planned for January 2012 (after the research in this thesis was completed).

5.11 Discussion and comparison to previous results

The statistical results are able to provide a description of the general F-region neutral atmosphere under the cusp; while the high/low geomagnetic and solar activity data provide windows to the northern/southern cusp boundaries. These are the first F-region neutral atmospheric cusp results provided by the FPI and SCANDI, and the first comparison of observations of emission intensities with neutral temperature and vertical wind for cusp investigations.

5.11.1 Comparison with CHAMP satellite data

The CHAMP satellite [Reigber et al., 2002] is able to detect density variations at altitudes of around 400km. Recent observations by Lühr et al. [2004] show neutral density values can be significantly enhanced by nearly a factor of 2 for the northern dayside cusp. Lühr et al. [2004] also suggest continuous upwelling lasting several hours at least is required for the density enhancement seen. Intense, small-scale, field-aligned current filaments are always seen with the density enhancement, and Joule heating is suggested as the primary cause of the thermospheric upwelling. The FPI statistical results demonstrate F-region neutral upwelling close to the centre of the cusp, reaching values of 100 m/s. Neutral temperature enhancements reaching values of above 1000K are also seen for solar maximum conditions. Joule heating and ion-frictional heating are the most common cause of neutral temperature

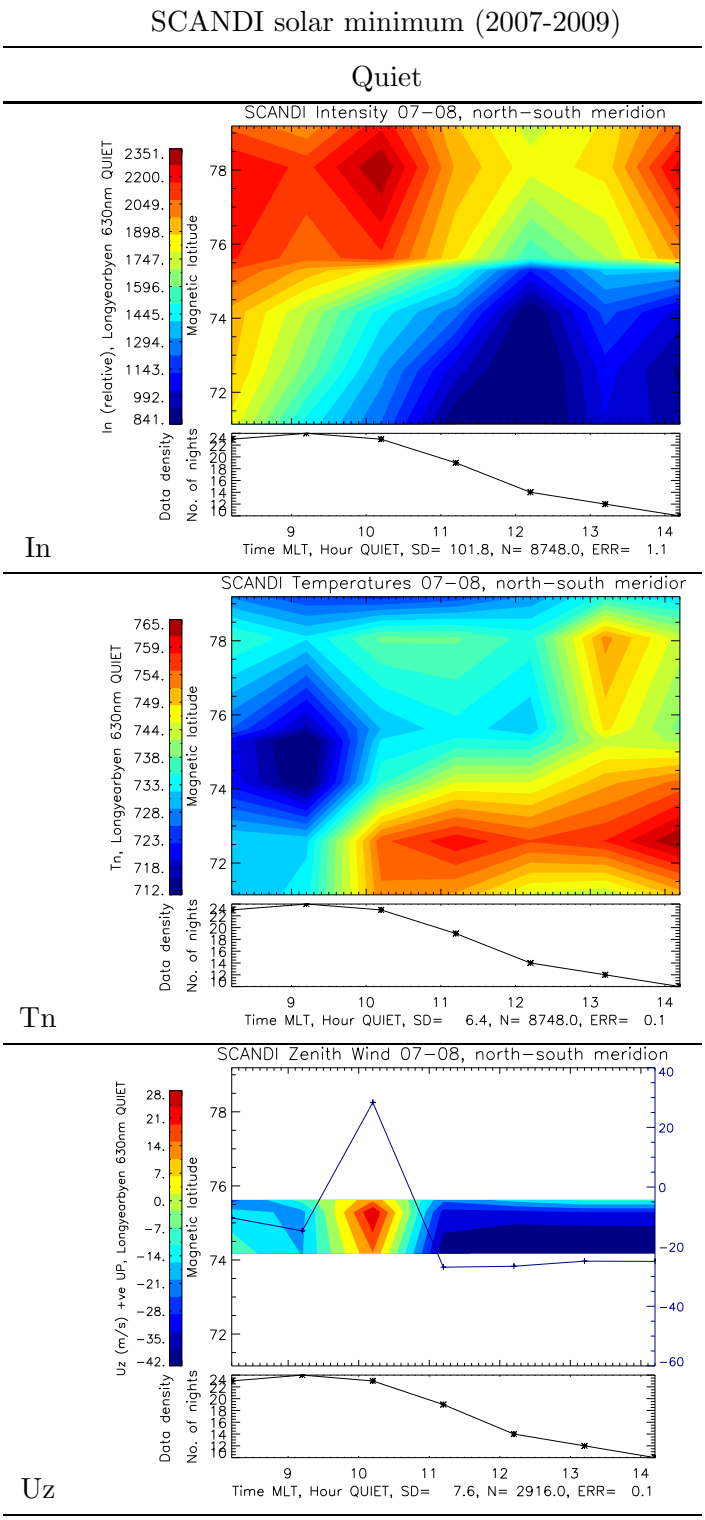


Figure 5.24: This figure is also similar to figure 5.22. However, SCANDI data with higher spatial resolution are shown, and they cover years 2007-2009, also at solar minimum. Only quiet activity data are shown because there were not enough moderate activity data available for this time period.

increases in the F-region. The FPI results here support the results from the Lühr et al. [2004] study.

The Schlegel et al. [2005] study involves a much larger data set, also from CHAMP. It is able to detect density troughs, as well as density peaks related to the cusp, with maximum amplitudes of 50%. The troughs tend to be broader and are found at higher latitudes than the peaks. The results can be compared to the FPI statistical results from the solar maximum-moderate activity case, where the most northern boundary of the cusp is observed within the FPI FOV. Temperatures reaching above 1000 K remain present in the F-region, and downwellings between -40 and -100 m/s are detected throughout the cusp time period. The FPI downwelling may account for the higher latitude density trough shown by the Schlegel et al. [2005] study, although this needs to be tested by model simulations. Overall, the FPI results here are complementary to the density effects seen by the CHAMP satellite.

5.11.2 Comparison with model data

The FPI results can also be further compared to the cusp-modelled results presented in Demars and Schunk [2007]. Demars and Schunk [2007] aim to simulate the cusp density enhancement seen by CHAMP by artificially increasing the ion-frictional heating between 60 and 70 degrees northern magnetic latitude and 10 and 11 MLT by a factor of 110-120. The authors acknowledge this is for a larger heating effect than occurs naturally.

Density, upward drift, wind vectors and neutral temperature results from Demars and Schunk [2007] are shown in figure 5.25. Density reaching 100% of the background value is simulated. Their model results the neutral gas moves upward within the heated cusp region, then spreads out both northward and southward outside of it to create a neutral 'fountain'. The northward (poleward) spread then follows the general anti-sunward neutral flow, while the southward (equatorward) spread remains sunward until 35 degrees co-latitude. The upward drift and neutral temperature model results are very similar to the FPI upwelling and temperature averages.

Comparing the model data with FPI results, the neutral temperature and upward drift results mostly match well with the FPI statistical data for neutral temperature and upwelling. Again, treating the FPI FOVs as covering the northern/southern sides of the cusp for solar maximum/minimum conditions (figures 5.22 and 5.23):

When the FPI is observing the northern side of the cusp at solar maximum, a strong

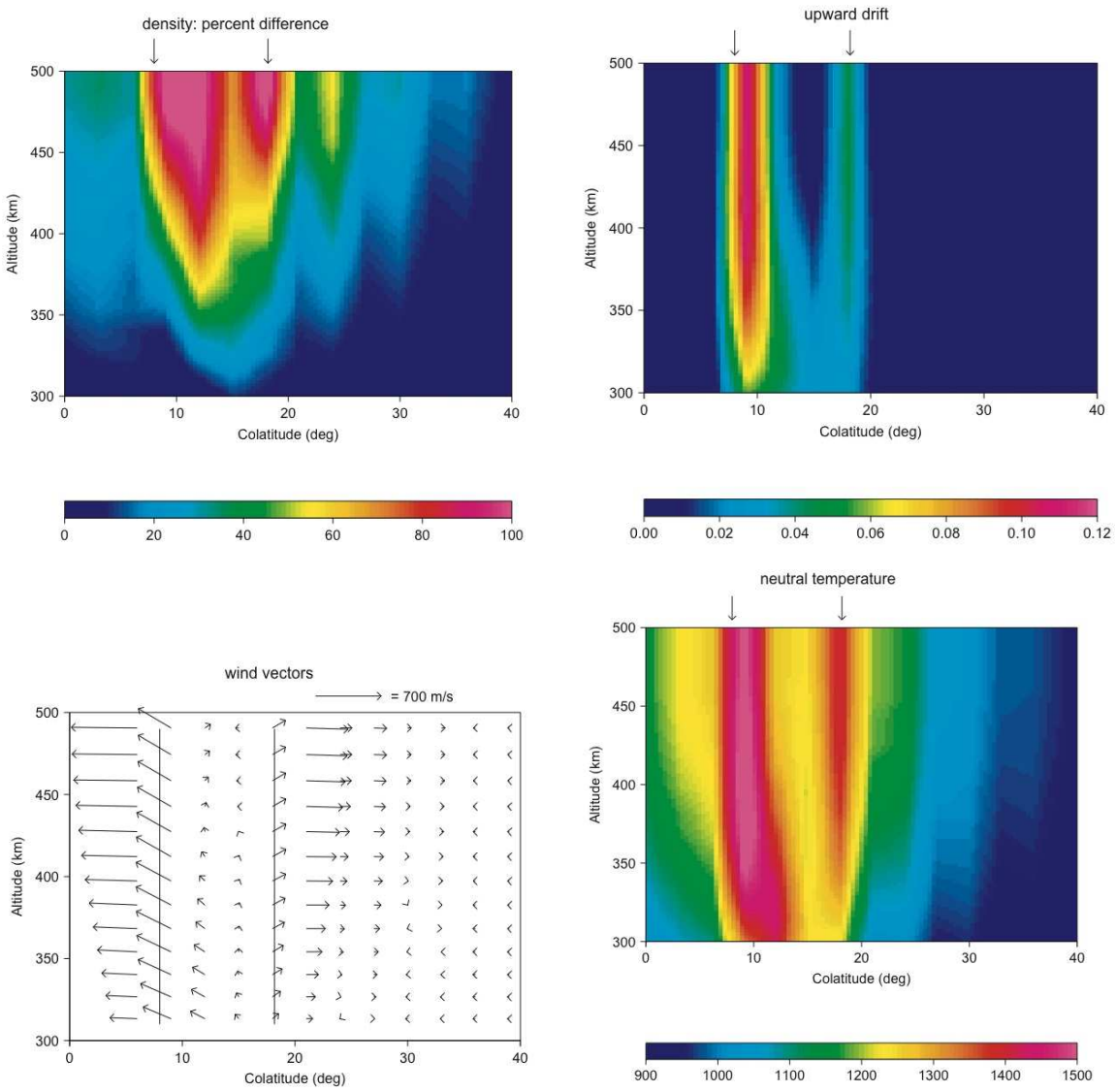


Figure 5.25: These plots are the Demars and Schunk [2007] cusp model results. The top row contains the density (left) and upward drift (right) and the bottom row contains the wind vectors (left) and neutral temperature (right).

upwelling reaching 100 m/s is seen under the maximum-quiet condition when the observation is closer to the cusp centre, as well as temperatures reaching above 1000 K. The highest temperature enhancements are found on the boundaries of the cusp in the maximum-moderate plot, similar to the model. However, downwelling is also seen at the cusp boundary in the FPI data.

When the FPI is observing the southern side of the cusp at solar minimum conditions, the highest temperature enhancement is again found at the cusp boundary (minimum-quiet), with smaller temperature increase detected overhead. Both upwelling and downwelling are detected at the boundary, but upwelling is the dominant effect when the FPI FOV is closer to the cusp centre (minimum-moderate).

The latitude locations of the enhanced boundaries detected by the FPI at solar maximum and minimum are also comparable with the model results, with consideration that the FPI FOV covers about 10 degrees latitude.

5.11.2.1 Density trough and downwelling

The only feature that is not repeated by the model is the downwelling seen by FPI when it is observing the most northern/southern boundaries of the cusp under maximum-moderate/ minimum-quiet conditions. The downwelling is presumed to be related to the density trough at the northern side of the cusp observed by Schlegel et al. [2005]. However, the Demars and Schunk [2007] model is also unable to simulate the density trough detected by Schlegel et al. [2005], and the model is therefore not expected to be able to accurately present the neutral zenith wind related to the density trough.

5.11.2.2 Cusp boundary effect

The reasoning for the enhancements in the boundary provided by Demars and Schunk [2007] is that in general, the ions are convecting across the simulated cusp region from its equatorward to its poleward sides. Therefore:

- At the equatorward boundary, the sunward flow of the neutrals, due to the fountain effect, is in the opposite direction to the antisunward ion convection, resulting in increased ion-neutral frictional heating.
- At the poleward boundary, the ion convection speed is large; therefore also contributes to increased ion-neutral frictional heating

However, Demars and Schunk [2007]'s explanation of the boundary heating cannot be supported by the FPI results in the current form. The FPI neutral wind clearly shows an antisunward flow, and not sunward flow, when it is observing the equatorward cusp boundary as shown in the solar minimum case study 3 (figure 5.21). An anti-sunward flow also cannot be clearly identified by the poleward boundary observation by case studies 1 and 2 at solar maximum (figures 5.5 and 5.11). Importantly, this could be because the model results are representing results for 300 km altitude and above, while the FPI is observing altitude between 210-260 km at a different altitude regime. Nevertheless, Joule heating and ion-neutral frictional heating as suggested by Lühr et al. [2004] should play an important role in the heating seen in the investigation here.

The more appropriate explanation of the intense rise in FPI temperature at the cusp boundaries is likely to be the theory provided in case study 1 (section 5.6.3.3). It demonstrates a westward neutral wind induced under cusp conditions through ion-neutral coupling. As the ion velocity subsides, the neutral wind remains westward because of neutral inertia. Large ion-neutral frictional and hence Joule heating is therefore induced at the cusp boundary.

5.11.3 Discussion of evidence for density trough and source of neutral heating

The satellite observes the atmosphere at a fixed height. Hence, a reduced/more-compressed atmosphere is associated with down-welling, and hence density troughs. A more compressed atmosphere is therefore associated with density troughs, related by the simple gas law (PV=nRT).

The three case studies in this chapter are for the purpose of providing evidence that the ion-neutral velocity difference and hence Joule heating is the cause of the increased neutral temperature at the cusp boundaries. Apart from Joule heating, the other possible cause of heating is precipitation from the cusp. Other possible sources of precipitation include the LLBL (equatorward of the cusp) and mantle (poleward of the cusp), as shown in figure 2.14. However, the electrons are more energetic, penetrate deeper and deposit energy lower down than the F-region and red-line emission altitude. Hence, they are not observed by the FPI. The conclusion is therefore supported by the case studies, as well as the statistical location of the different types of precipitation at high latitude site.

5.12 Summary and Conclusion

The three case studies present results from three experiments representing different geomagnetic and solar conditions; with the first two case studies showing results from solar maximum, geomagnetically active periods; and the third case study showing results from solar minimum, quiet period. They demonstrate, as individual cases, that the cusp is frequently within the FPI FOV between 8-14 MLT. Temperature increases are seen in regions directly at the cusp, as well as at the boundaries, and upwelling is also detected. For the solar maximum FPI case studies, a strong increase in the zonal westward flow is associated with the cusp, but the effect on the quiet case study is less clear.

The statistical average FPI results are able to present FPI observations on the northern boundary of the cusp, when the auroral oval is expanded equatorward under solar maximum and geomagnetically active conditions. The FPI also observes the southern boundary of the cusp when the cusp is contracted with the auroral oval at solar minimum and quiet conditions. Overall, the statistical results are supported by the individual case studies. Their comparisons with CHAMP satellite density data are also complementary to each other, with temperature enhancements, as well as upwelling and downwelling locations matching well with the density peaks and troughs observed overhead and at the northern side of the cusp. Comparison with model results further confirms the temperature and upwelling enhancements at the cusp boundaries, although the model does not replicate the downwelling region. Overall, the F-region temperature and upwelling enhancements are likely to be linked to ion-neutral frictional and Joule heating.

The FPI and SCANDI results here present the first directly measured neutral parameters over the cusp region. It is further supported by 3 case studies. The statistical data is able to provide the general neutral conditions on the southern and northern sides of the cusps, which are results that have not been presented before. The comparison of FPI data with model also indicates the current method of simulating cusp dynamics and density requires improvement. It points to the importance of a higher horizontal spatial resolution thermospheric model to allow for larger than expected but localised vertical winds, and suitable levels of heating to match observations.

Chapter 6

Summary and conclusion

6.1 Overview

This thesis describes the development of a new optical instrument, SCANDI. It provides background on the instrumentation hardware, as well as the data analysis process. The advance in temporal and spatial resolution for SCANDI over the single-view FPI has been highlighted. Data calibration and quality checks are shown to demonstrate how the quality of SCANDI data can be maintained. The thesis then applies both FPI and SCANDI statistical data to a temperature trend study. This study devises a method for the application of directly measured FPI temperatures in temperature trend investigations, with consideration to the airglow profile. The importance of the effect of changes in emission height on the neutral data are presented and temperature data from the last ten years are shown. The data are then compared against MSIS model results and their differences are analysed. The causes of these differences are investigated with correlation tests using geomagnetic indices. The results demonstrate that MSIS underestimates the neutral temperature by 200K in this usually weak solar cycle at solar maximum. This is the longest neutral temperature investigation that has been presented at the time of this thesis, and it enables the distinction of solar maximum and minimum data, which has not been done before. The diurnal variation of neutral temperature data are compared with CMAT2 results. This study utilises the high spatial resolution data from SCANDI and enables comparison between the instrument and model over a latitudinal range. It shows a reasonable spatial match, but the magnitudes presented by CMAT2 are smaller. Finally, the spatial resolution of SCANDI is again utilised for a cusp study from the neutral atmosphere's point of view. A statistical study is presented which shows direct

neutral measurements that support the conclusion of a CHAMP cusp density investigation. Similarly, it also supports previous cusp neutral model results. The statistical analysis is also accompanied by three case studies with comparisons of EISCAT, MSP and FPI/ SCANDI data.

6.2 Instrumentation

The main purpose of the instrumentation chapter is to document the new UCL instrument, which is an all-sky version of the Fabry-Perot Interferometer (FPI), named the SCANning Doppler Imager (SCANDI). This includes the basic description of the instrumentation hardware and control software. The data analysis and calibration methods are also shown, as well as examples of SCANDI results.

In terms of SCANDI instrumentation, particular attention is paid to the all-sky imaging system, the capacitance stabilised etalon and the electron multiplying CCD imager. They are new techniques that are applied to SCANDI. The phase compensation method used in a scanned and imaging etalon is the main focus in the control software description. The method for adapting the instrument to observe different emission lines are also provided.

The data analysis and calibration section includes a description of the fish-eye lens projection, with consideration to the different types of projection methods. It is decided that a simple least-square fit relating the incoming rays' zenith angle and the final CCD pixel position is the most suitable method. Furthermore, the physical position of the instrumental optical axis on the EMCCD detector is deduced. The equivalent emission volume observed by each SCANDI zone is also considered for both flat and curved Earth projections. It is decided that the flat projection is suitable for the 25-zone resolution used in SCANDI currently (up to observing season 2010/11). It gives an approximately equal volume of the detected emission layer for each zone. However, when the spatial resolution increases, the effect of the curvature of the Earth becomes more significant, especially for the outer SCANDI ring. It could cause a difference of $\sim$27%. The high resolution SCANDI projection is designed and ready for future applications. It can be applied when higher spatial resolution measurements become possible for the more active geomagnetic conditions, as we move away from the extended solar minimum (cycle 23/24).

The data reduction and retrieval process which produces the emission intensity, neutral temperature and line-of-sight winds are then provided. The error estimation method for the parameters are also described. This includes the calibration work. It involves the fine-

tuning of the etalon parallelism during a scan, which is able to correct the temperature ring-offsets seen in the earlier temperature data. The central zone temperature offset is found to be caused by the sensitivity of the fitting program to initial parameters, and is corrected. The different methods of deducing the line-of-sight wind baseline/zero Doppler shift position are investigated, which show that the baseline follows a sinusoidal variation over consecutive azimuth angles for zones in a SCANDI ring. It is decided that the cause is likely to be instrumentational, and needs to be corrected when deducing the zero Doppler shift position. The general approach is to determine a separate baseline for each zone from the zone's daily average. This is suitable for the immediate analysis during an observing season. The data are then fine-tuned with the longer-term sine-baseline which requires averaging over more than one night of data. The determination of the appropriate baseline is further tested by direct comparisons and correlation test, which apply the LOS wind measurements between the co-located SCANDI and FPI.

The horizontal wind fitting method is discussed and verified, which requires the assumption of uniform vertical wind field. The original method is to apply the zenith wind measurements from the central zone, and assumes it can represent the vertical wind over the whole field-of-view. This is usually valid since the vertical wind is relatively small comparing to the horizontal wind. However, when a localised large zenith wind is detected by the zenith zone, it is unrealistic to assume the whole field-of-view has the same zenith wind component. The conclusion is that it is best to assume the vertical wind is negligible comparing to the horizontal wind field in the wind field fitting process, since large vertical wind cannot be accounted for correctly by the current method.

Finally, smaller projects including a fixed-gap etalon analysis and examples of some typical SCANDI results are shown.

6.3 FPI neutral temperature trend study in the polar region

This chapter first establishes the best method for collating FPI and SCANDI data in a long term temperature trend study. It shows that under cloudy conditions, the temperature varies linearly with the magnitude of neutral wind. Since neutral wind has a diurnal pattern, the full 24-hour daily average temperature result is able to average out the cloudy effect. It shows that the 24-hour daily average temperature can agree within 95% confidence level even if cloudy data are included. This can significantly improve the FPI data sample size in a long-term temperature trend study. However, cloudy data for samples

that are smaller than a 24-hour period need to be excluded.

The chapter than applies existing airglow models in FPI/SCANDI studies. The FPIs measure the neutral atmosphere through red and green-line emissions for the F and E-regions respectively. Traditionally, the FPI measurements are assumed to be taken at a fixed altitude at the peak of the emission profile. However, the actual measurements are weighted by the emission profile, and are therefore subjected to variations in the peak emission altitude, as well as its profile shape. In terms of the peak emission altitude, there is negligible difference between some of the airglow models that are mentioned in the thesis.

The red line peak emission altitude is further compared against hmF2 altitude, and is able to match within 15km (about 1.5 scale height) to the peak emission altitude determined for the model. It is also demonstrated that the peak emission altitude can change at scales of 10s of km within 10s of minutes. It is shown that ignoring the altitude variation in ion and neutral temperature comparisons could lead to a discrepancy of around 50K. The discrepancy is reduced to 10K if the correct altitude is applied. This highlights the importance of the consideration of emission height variation in FPI measurement for event-based studies.

The neutral temperature solar cycle variation for cycle 23/24 is then investigated, which covers years 1999-2009. Data sets that have full 24-hour darkness at the Svalbard site are chosen. This covers November, December and January of each year.

For the solar maximum period between 2001-2004, the average FPI-measured neutral temperature shows an extended, but unexpectedly lower than the average MSIS-model neutral temperature value. The measurement is lower than modelled by 200K. The lower than modelled temperature is found at quiet solar and geomagnetic activity periods. This is supported by the contour plots (T_n vs. Ap vs. $F_{10.7}$) of the T_n values and densities. This indicates that MSIS overestimates the low activity temperature at solar maximum, possibly because this solar cycle is unusually weak. The wider red line emission profile measured by the FPI may also contribute to the lower than modelled temperature. In terms of correlation, the neutral temperature is better correlated with $F_{10.7}$ than Kp indices at solar maximum. However, the correlation remains low in terms of the daily average results.

For the solar minimum period between 2008 and 2009, the average neutral temperature is higher than the modelled results by approximately 100K, with consideration to the

altitude variations of the red line emission peak. This is expected because MSIS is known to underestimate temperatures at high activity periods in general. The reported low mass density in the atmosphere may also affect the accuracy of the temperature profile modelled by MSIS.

Furthermore, the average solar minimum temperature is around 750K, which is similar to the reported temperature from previous solar cycles. This appears to contradict the lower than normal global temperature reported by other authors for the extended solar minimum. One possibility is that the polar cap site may have higher than expected heating from a contracted oval and geomagnetic sources. In terms of correlation, the neutral temperature is better correlated with Kp indices than $F_{10.7}$ at solar minimum, although correlation also remains low for the daily average results.

The correlation test is then applied on hourly average results, with the consideration to neutral atmospheric inertia and the related time delay in neutral atmospheric response. The highest correlation of 0.83 between the hourly temperature and Kp is found when the neutral temperature is delayed by 5 hours. The time delay matches with the values found in previous atmospheric inertia research. The results are applicable for solar maximum data only. The solar minimum hourly data correlation remains low. The new conclusion here is important, where previous reports are not able separate the solar maximum and minimum data because of their shorter (half a solar cycle) data sets. Atmospheric inertia are also not considered by previous statistical temperature investigations.

Finally, the different possible solar indices including AE, SEM and $S_{10.7}$ are correlated against the daily average T_n data. They show negligible difference from the correlation results of $F_{10.7}$ and Kp. Atmospheric models such as MSIS and CMAT2 apply these indices as a guideline for their temperature results. The investigation here indicates that using different indices in an atmospheric model has insignificant effect on the modelling of polar cap, daily-averaged, neutral temperatures.

The diurnal variation of neutral temperature is then investigated, which shows the expected variation due to changes in horizontal advection and particle precipitation heating. These are the two dominant heating sources. The main diurnal peak is found at 21-25MLT, with a diurnal variation of ~600K. This is significantly greater than the diurnal variation predicted by MSIS. A smaller peak is also detected at 9-14MLT relative to the cusp. The different heating terms responsible for the diurnal pattern are also modelled by CMAT2, although it is unable to match the main diurnal peak, and gives a modelled

peak that is lower by ∼150K.

6.4 The cusp and the neutral atmospheric in the F-region

The neutral thermospheric view of the cusp is investigated. The solar-cycle length FPI data set, as well as high spatial resolution SCANDI data from both the E and F-regions are utilised.

In order to support the statistical results, three case studies representing different geomagnetic conditions are included. Each case study presents a comparison of EISCAT ESR, MSP and FPI/SCANDI data. In these case studies, the ionospheric cusp dynamics are already well researched by previous experiments. The additional neutral data therefore add to the understanding of the cusp dynamics. Overall, the case studies support what is being shown by the statistical study in this chapter.

For the statistical study, the FPI and SCANDI data are able to show that the cusp is located at lower latitudes at solar maximum, and higher latitudes at solar minimum. This is directly related to the auroral oval and magnetopause position variation over the solar cycle.

The FPI data also in general complements the CHAMP satellite density data, with the neutral temperature enhancements, upwelling and downwelling locations matching well with density peaks and troughs. Comparisons with model results also re-confirm the upwelling and temperature enhancement regions at the cusp boundary. This indicates the link of cusp upwelling mechanisms with F-region ion-neutral frictional and Joule heating. This is a new study which applies direct measurements of the neutral atmosphere to cusp observations.

6.5 Future work

The temporal and spatial scale sizes for the auroral atmosphere have been challenged by our FPI and SCANDI observations. The cusp study in this thesis is one of the examples. Lower integration time and higher spatial resolution can both be achieved with continued instrumentation developments such as CCD sensitivity. Therefore, both the number of SCANDI zones and time resolution can be optimised further in the future. For example, at the time of this thesis (2005-2011), SCANDI was operating using only a division of 25 zones , with a time resolution of approximately 5 minutes, which was appropriate for

the low solar flux levels of the extended solar minimum. As of observing season 2011, in January 2012 the resolution was increased to 61 zones in response to increased solar flux, and then for season 2012, the resolution was increased further to 91 zones with a similar time resolution.

Both SCANDI and the FPI are located at Longyearbyen, Svalbard. It is a unique observing site with instruments such as the EISCAT Svalbard Radar and a number of auroral imagers. The combinations of instruments with the FPIs are able to provide simultaneous observations of the neutral thermosphere and the ionosphere. Multi-instrument observation is essential for the complex dayside cusp studies. One example of an application of work presented in this thesis are the cusp upwelling campaigns undertaken in January 2012 and January 2013. Both campaigns have observed large vertical winds (up to 200 m/s) in the cusp that last for an hour or more [Aruliah et al., 2013] . The results are contrary to the standard belief that vertical winds must be small. The results of this thesis contributed to the science and design of these campaigns and support the observations of small scale vertical winds.

The ten year study of neutral temperature provides the statistical temperature trend over a solar cycle. Although it is one of the longest continual polar thermospheric studies from a single site, it is not yet long enough to provide an indication of whether there is a longer-term cooling effect in the upper thermosphere. The comparison of $F_{10.7}$ and FPI temperature here is also currently unable to provide a statistically significant correlation and a clear trend beyond the 10 years. However, the methodology and selection criteria used for the data analysis, and the calibration techniques, are essential for longer term FPI temperature studies. One example of methodology is the incorporation of the emission model to account for emission height variations in this thesis. The extension of this study is to apply FPI temperatures from further into the future and further into the past to assess if any long term cooling effect is observed by the FPIs in the auroral zone. Analysing past FPI data would require developing a technique to derive temperatures when there was no laser to determine the instrument function. Another application of statistical FPI temperature data is the study of possible effects of sudden stratospheric warming (SSW) on the thermosphere. SSW is mostly seen in the northern winter polar region, where there is evidence that the SSW has influence on the thermosphere [Goncharenko et al., 2013].

Furthermore, Space Weather prediction requires better thermospheric modelling, in addition to the modelling of the ionosphere-magnetosphere coupling process. After all,

the thermosphere modulates the ionosphere-magnetosphere coupling. In order to achieve this, higher resolution modelling is required for the auroral region. High spatial and temporal resolution data from SCANDI and FPIs are therefore essential to achieve such developments in the GCMs.

Appendix A

Ionospheric background

The high latitude ionosphere is one of the major topics in this thesis, and is directly coupled to the neutral thermosphere that the FPI measures.

A.1 Ionospheric structure

The ionosphere is the ionised part of the atmosphere, which is formed by the ionisation of major neutral species (section 2.2.6) by EUV and X-ray solar radiation in the middle and low latitude regions, and additionally from collisions with energetic particles, especially for the high latitude areas. The main regions of the ionosphere are summarised in figure A.1.

The ionosphere is the part of the atmosphere that contains a significant number of free electrons and positive ions, and is electrically neutral. Hence the electron density is an important parameter for understanding the properties of the ionosphere. The electron density is governed by the rate of production and loss processes, and at dynamic equilibrium, the rate of change of electron density can be described by the continuity equation:

$$\frac{\partial N}{\partial t} = q - L - div(N\mathbf{v}) \tag{A.1}$$

where the three terms on the right are simply the production rate, loss rate and the loss of electrons by movement respectively. The main production processes are photo-ionisations from EUV and X-ray radiation, and from energetic particle ionisations. The final term of the continuity equation accounts for changes of electron and ion densities that are caused by the plasma bulk movement. The production and loss process terms in the equation are discussed further below.

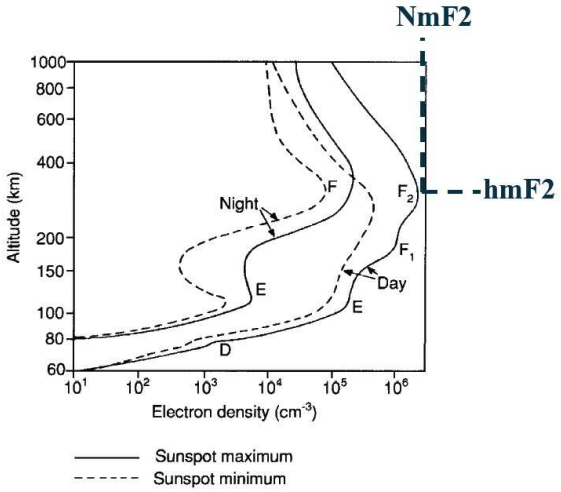


Figure A.1: This figure shows the main regions of the ionosphere, which are the D region at 60-90km, with density at 10^2-10^4cm^{-3}; the E-region at 105-160k with density at 10^5cm^{-3}; the F1 region at 160-180km with density 10^5-10^6cm^{-3}; and the F2 region at a maximum altitude of 300km, with density up to 10^6cm^{-3}. hmF2 and NmF2 denote the maximum height and density of the F2 layer respectively. Note that the D and F1 regions completely disappear at night times.

A.2 Ionosphere: production

The most fundamental description of the photo-ionisation production rate is through the Chapman production function, which is based on the simple rate of production of ion-electron pairs expressed as

$$q = \eta\sigma nI \tag{A.2}$$

where η is the ionisation efficiency, σ the absorption cross-section, and σnI is the total energy absorbed per unit volume of the atmosphere per unit time. I is the flux of incident radiation with units J/m^{2}s, and n is the density of molecules/ions that can be ionised at the radiation's wavelength. The derivation of the Chapman production function can be outlined by considering the intensity variation with distance, x, and applying equation A.2:

$$\frac{dI}{dx} = -\sigma nI \tag{A.3}$$

which can be written as

$$I = I_\infty e^{-\tau} \tag{A.4}$$

where I_∞ is the intensity outside of the atmosphere. τ is the optical depth, i.e. σN_T where N_T is the number of absorbing atoms/molecules in a unit column. N_T can be

written as nH.secχ, where H is the scale height, and χ is the solar zenith angle. Hence the optical depth is

$$\tau = \sigma(nH.sec\chi). \tag{A.5}$$

Finally applying the exponential atmosphere, $n=n_0e^{-z}$, with equation 2.5, where n_0 is the particle density at height z=0, the Chapman production rate is

$$q \propto e^{-z} \cdot e^{-K} exp(-z) \tag{A.6}$$

where K is a constant, and integrating this gives maximum q at $Ke^{-z}=1$, which is an useful result that shows the production rate is the greatest at the height where the optical depth (equation A.5) is unity, which directly determines at what altitude the ionosphere is found. This is an important factor in ionospheric studies because observations are usually taken at the maximum of an ionospheric layer. Note the actual atmosphere is more complicated than that described by the Chapman function, but it still provides a useful guideline in ionospheric experiments.

Another important production source is the ionisation from energetic particles, which are listed here. It is an important source at the high latitude regions, and the sources are:

- **Electrons.** The range for which a stream of electrons can travel depends on the rate of energy loss due to their collision, ionisation and excitation of neutral particles, and hence depends on atmospheric density and pressure. In a neutral gas, each ion-electron pair ionisation requires ∼35eV of energy.

- **Protons.** They are particularly important at high latitudes when solar flares cause fluxes of protons. Protons are more energetic than electrons, and can leave ionised trails as they collide with atmospheric gasses.

- **Bremsstrahlung X-rays.** They are the electromagnetic energy, in the X-ray range, that is released as the energetic electrons collide with neutral gases. The process is called Bremsstrahlung, and is the 'braking radiation' that is released as the electron decelerates quickly.

A.3 Ionosphere: loss

Two main types of loss processes are considered and they describe the general processes involve in chemical recombination, which is the major loss process in the ionosphere.

- If the electrons directly recombine with positive ions, and there are no negative ions present, the main electron loss process is simply X^++e→X, where X^+ is the positive ion. However, it should be noted that negative ions cannot be neglected for the lower atmosphere, and recombination can occur between positive and negative ions.
- An electron can also simply attach itself to a neutral particle, forming a negative ion i.e. M+e→M^-.
- In the upper atmosphere, especially for the F region, the electron loss appears in two stages. First, a molecule breaks apart and one part attaches to an ion X^++A_2 →AX^++A, and then the new charged molecule AX^+ recombines with an electron through AX^++e→A+X. This is dissociative recombination.

A.4 Ionosphere: outline of chemical processes

The E and F1 regions are at 105-110km and 160-180km altitudes respectively, and the main reactions in both the E and F regions are dissociative recombination. The major specific reactions are

$$e + O_2^+ \rightarrow O + O \tag{A.7}$$

$$e + N_2^+ \rightarrow N + N \tag{A.8}$$

$$e + NO^+ \rightarrow N + O. \tag{A.9}$$

Radiative recombination, e+$X^+ \rightarrow X + h\nu$, is not important in these regions because it is 10^5 slower than the major dissociative recombination reactions.

For the E region, 80.0-10.27 nm is part of the solar spectrum that is absorbed in the EUV, and is mostly absorbed by molecular oxygen to form O_2^+. X-rays (wavelength 1-10 nm) also play a part in the E-region, and can ionise all of the neutral species in this part of the atmosphere, producing mainly NO^+, O_2^+ and O^+. However, X-rays have limited effect at solar minima, because their intensity is directly linked with the solar cycle variations.

For the F1 region, 20.0-90.0nm is the part of the spectrum that is absorbed, which is also in the EUV. The major species produced are NO^+ and O_2^+ at the end result of the overall chemical reactions, though the primary reactions also produce intermediate species, and they are O_2^+, N_2^+, O^+, He^+ and N^+. In the F region, the reactions that are important are the charge exchange processes. Using O_2 reaction as an example, and similar reactions can be applied to N_2:

$$O^+ + O_2 \rightarrow O_2^+ + O \tag{A.10}$$

$$e + O^+ + O_2 \rightarrow O + O + O. \tag{A.11}$$

Note that the F1 layer disappears at night, and the E layer becomes weaker.

The F2 layer is one of the most important layers for the results in this thesis, since most analysis from this thesis focuses on the red-line emission in the F layer. The F2 layer peaks at 200-400km. Its mechanism is different from those of the F1 and E regions, where none of the wavelengths of solar radiation reaches maximum ionisation in the F2 region. The ion production rate simply decrease with altitude. The F2 layer is explained with the variations in the recombination process, which involves the two stage recombination processes as follows:

$$O^+ + N_2 \rightarrow NO^+ + N \tag{A.12}$$

$$e + NO^+ \rightarrow N + O \tag{A.13}$$

taking O^+ as the major species in the F2 region as an example: The recombination rate depends on the concentration of N_2, associated with reaction A.12, while the production rate depends on the concentration of O. Hence the O to N_2 ratio is an important factor in the F2 layer. This reaction produces an electron density variation that increases with height because the loss rate decreases more quickly than the production rate. Diffusion also becomes important at F2 altitude because of the low air density. Hence, the F2 layer peaks at the point where the two processes above become equally important.

Finally, emission of radiation, i.e. airglow and aurora, are important reactions that appear with the chemical processes discussed here. The airglow is constantly present in the atmosphere, while aurora is present at high latitude regions, and appears during and after geomagnetic storms. In general, aurora and airglow are caused by photon emissions from radiative recombination, and reactions with excited states caused by electron collisions,

electric field and solar radiations. The energy levels of the green and red-line emissions, which are the two major reactions observed by the FPIs in this thesis, are shown in figure A.2.

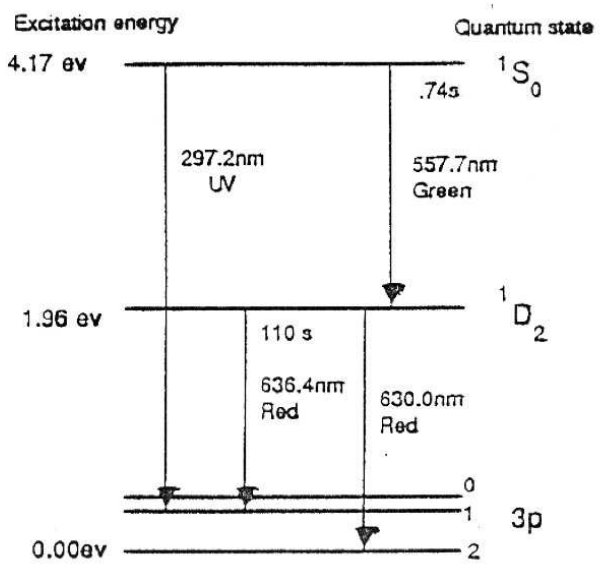


Figure A.2: Courtesy Hargreaves [1992]. This figure shows the major transition lines, including the green-line (557.7nm) and red-line (630.0nm) transmissions.

Appendix B

Conductivity

The conductivity is

$$\sigma = \frac{i}{E} = \frac{Nev}{E} \tag{B.1}$$

where i is the current density in Am^{-2}, N is the density of the particles, e is the electric charge, and v is the particle velocity.

Consider the situation where the electric field components are perpendicular to the magnetic field, which is similar to the methods in the previous velocity derivations, but the electric field may have both an x and y component. Applying the definition of conductivity (equation 2.15), the electron and ion velocities as shown in table 2.1, and replaces the general force, F, with electric field, Ee, for ions and -Ee for electrons, gives the Pederson conductivity:

$$\sigma_1 = \frac{Nev}{E} = \frac{e}{E} \cdot (N_i v_i + N_e v_e) = \frac{e}{E} \cdot \frac{E}{B}(N_i f_1(i) + N_e f_1(e)) \tag{B.2}$$

which with ω_B =eB/m can be expressed as

$$\sigma_1 = e^2 \left[\frac{N_e}{m_e \nu_e} \cdot \frac{\nu_e^2}{(\nu_e^2 + \omega_e^2)} + \frac{N_i}{m_i \nu_i} \cdot \frac{\nu_i^2}{(\nu_i^2 + \omega_i^2)} \right]. \tag{B.3}$$

Similarly for the Hall conductivity:

$$\sigma_2 = \frac{Nev}{E} = \frac{e}{E} \cdot (-N_i v_i + N_e v_e) = \frac{e}{E} \cdot \frac{E}{B}(-N_i f_2(i) + N_e f_2(e)) \tag{B.4}$$

$$\sigma_1 = e^2 \left[\frac{N_e}{m_e \nu_e} \cdot \frac{\omega_e \nu_e}{(\nu_e^2 + \omega_e^2)} - \frac{N_i}{m_i \nu_i} \cdot \frac{\omega_i \nu_i}{(\nu_i^2 + \omega_i^2)} \right]. \tag{B.5}$$

Finally, the conductivity parallel to the magnetic field is

$$\sigma_0 = \left(\frac{N_e}{m_e \nu_e} + \frac{N_i}{m_i \nu_i} \right). \tag{B.6}$$

The current density in tensor form, with the electric field also in factor form can be written as

$$\begin{pmatrix} J_x \\ J_y \\ J_z \end{pmatrix} = \begin{pmatrix} \sigma_1 & \sigma_2 & 0 \\ -\sigma_2 & \sigma_1 & 0 \\ 0 & 0 & \sigma_0 \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix}. \tag{B.7}$$

Appendix C

European Incoherent SCATter (EISCAT) radar

The advantages of incoherent scatter radars are that they are not restricted to the region below (or above) the altitude of peak electron density such as, for example, the ionosondes, and can observe both sides of the peak simultaneously. They can also achieve higher spatial resolution because the antenna has to be large relative to the radio wavelength. The disadvantage is that very weak signals are detected, which require a sensitive receiver and advanced data processing techniques.

The physical basis of the incoherent scatter radar is Thomson scatter, which is the re-radiation of incident electromagnetic energy by free electrons. The total 'radar cross-section', which is defined in terms of spherical scattering in all directions equally for a typical scattering region (with 10^{12} electrons/m^3 and a volume of $1.4x10^{-5}m^3$) is only equivalent to a sphere of ~2mm radius. Hence a radar must be extremely sensitive.

The radar is expected to detect a spectrum of Doppler shifts, produced by electrons moving at thermal velocities in the medium. However, the spectrum detected by the radar is ~200 times narrower than expected from simple thermal electron movements. This is because the electron movements are controlled by the ions.

The extent to which the ions control electron motions depends on the radar wavelength relative to the Debye length, λ_D i.e. the measure of the sphere of influence of a test charge placed in a plasma:

$$\lambda_D = \sqrt{\frac{\varepsilon_0 k T_e}{N e^2}} \tag{C.1}$$

where ε_0 is the permittivity electric constant, k is the Boltzmann's constant, T_e the electron temperature, N the plasma density and e is one electron charge. The ionospheric λ_D can range from a few mm in the F-region to about a cm in the D-region.

If the radar radio wavelength, λ is smaller than λ_D, the electron motions detected are controlled by electron temperature and mass; and if λ is greater than λ_D, the scattering is mainly due to irregularities of electron density and controlled by the ions.

It can be approximated that the radar detection is an assembly of irregularities, and if the collective behaviour of the plasma has a wavelength greater than the Debye length, the radar is able to detect them. The Debye length, moving at the ionic thermal speed $v_i = \sqrt{2kT_i/m_i}$. Hence the Doppler shift (2v/λ) can be written as

$$\Delta f_i = \frac{(8kT_i/m_i)^{\frac{1}{2}}}{\lambda}. \tag{C.2}$$

A typical Δf_i is ∼3500 Hz for pure O^+ ions at T_i=1600K and λ=75cm. Also if $\lambda < \lambda_D$, the electron line is detected instead because the scattering mainly comes from electrons that are not controlled by the ions.

In general, both electron and ion lines are present in the incoherent scatter spectrum. A typical return ion spectrum is shown in figure C.1. Though not displayed in the diagram, the electron plasma line is also usually measured by the radar at frequencies of ∼MHz, while the ion spectra centres around the ∼kHz region. For a cold unmagnetised plasma, the natural frequency at which electrons will oscillate is described by

$$2\pi f_p = \omega_p = \sqrt{\frac{N_e e^2}{m_e \varepsilon_0}} \tag{C.3}$$

where N_e is the electron density and m_e is the electron mass.

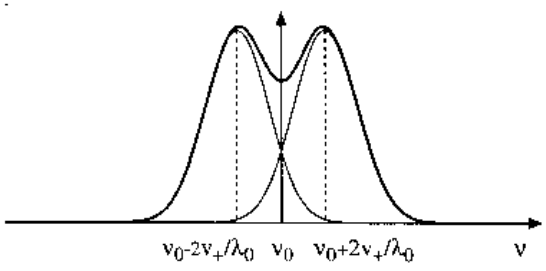


Figure C.1: This shows a typical return spectrum from incoherent scatter. The typical width of the spectrum is around a few kHz. Though not shown here, the electron/plasma lines can also be found at +/- MHz.

A more exact treatment of the spectrum, which considers the fluctuations within the medium as a spectrum of ion-acoustic and electron acoustic waves, is used by EISCAT.

The following is a list of some of the parameters that can be deduced from the spectra:

- Plasma velocity i.e. the bulk plasma drift in the radar's LOS, from the Doppler shift of the whole spectrum.
- Electron density, from the total power returned, which does not give absolute values of electron density, and requires calibration against other instruments such as an ionosonde. Another Ne calibration source comes from the plasma lines, though they are not always strong enough to be detected.
- T_e/T_i ratio, from the ratio of peaks to dip in the spectrum.
- T_i/m_i ratio, from the separation of the peaks. The usual method is to model the ion density from different species, deduce m_i, and hence T_i. In theory, it is also possible to identify transitions of some of the ion species, and hence their densities, since the detailed shape of the spectrum is affected by different mixtures of ions.
- T_e and T_i, from the line-width, if m_i is known as described above.

Some of the effects of changes in these parameters on the returned spectrum are shown in figure C.2.

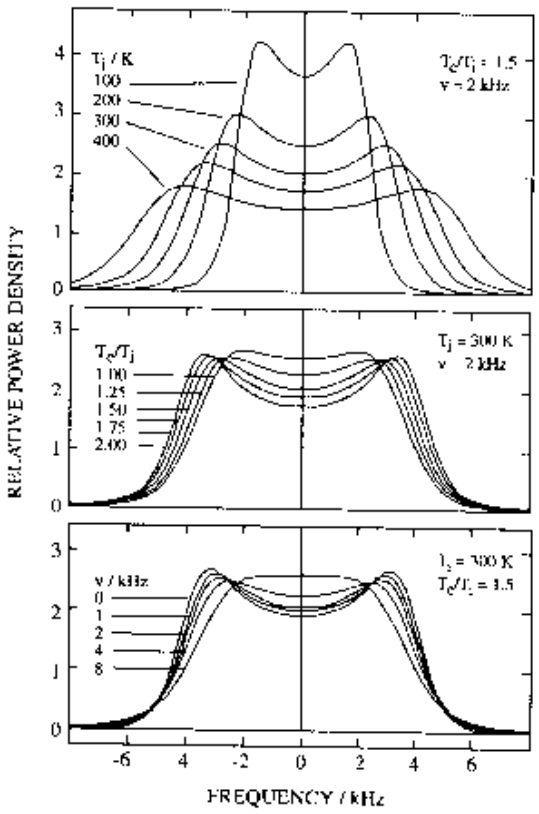


Figure C.2: This shows a typical return spectra from incoherent scatter. The top diagram shows the effect of changing T_i on the spectrum at fixed T_e/T_i and bulk ion frequency, the second diagram shows the effect of changing T_e/T_i at fixed T_i and ion frequency, and bottom diagram shows the effect of changing bulk ion frequency, and fixed T_i and T_e/T_i.

Appendix D

Determination of SCANDI zone sizes and solid angles

This appendix contains the methods used for determining the zone sizes and solid angles of the SCANDI zones when they are projected onto the red-line altitude, at approximately 240km. Earth is assumed to be a perfect sphere, and the emission layer is a simple spherical shell surrounding Earth's surface. Examples of SCANDI observing areas in terms of 'zones' and 'rings' can be found in e.g. figures 3.12 and 3.13.

D.1 SCANDI zone sizes, with considerations of the curvature of the Earth

As demonstrated by figure D.1, and by simple Pythagorean theorem:

$$(R - h)^2 + a^2 = R^2 \tag{D.1}$$

where R=h_e+h_s is the distance from centre of the Earth to the emission layer. h_e is Earth's radius, and h_s is the height of the emission layer from Earth's surface. Also by simple geometry, a, the radius of the spherical cap is

$$a = (h_s - h)\ tan(\theta_s) \tag{D.2}$$

where h is the height of the spherical cap, and θ_s is the zenith angle of a SCANDI zone from the SCANDI observing centre on Earth's surface. Hence, combining equations D.1 and D.2, and rearranging to solve for h gives the follow quadratic coefficients:

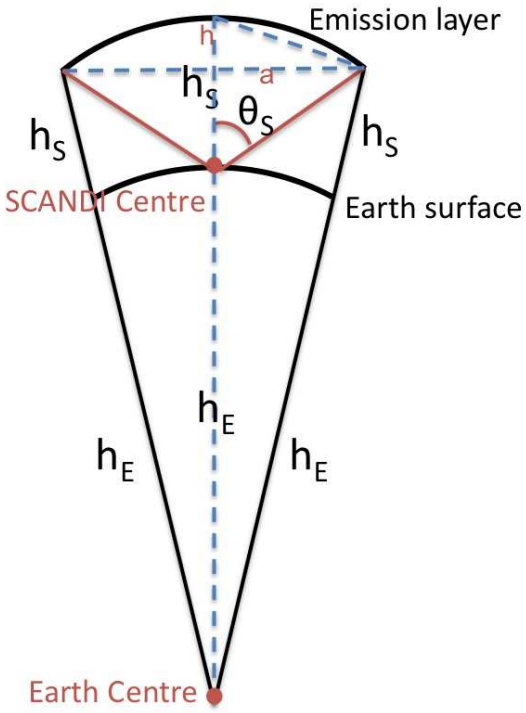


Figure D.1: This figure demonstrates the relationship between SCANDI's observing centre, Earth's centre, the emission layer and Earth's surface. Hence, a simple spherical cap area equation can be used to calculate the zone sizes. θ_s is the zenith angle of a SCANDI observing zone, h_s is the height of the emission layer determined from the SCANDI centre on the surface of the Earth, h_e is the radius of the Earth, a is the radius of the spherical cap, and h is the height of the cap.

$$a_q = 1 + tan^2\theta_s \tag{D.3}$$

$$b_q = -(2R + 2h_s tan^2\theta_s) \tag{D.4}$$

$$c_q = h_s^2 tan^2\theta_s \tag{D.5}$$

and the following root of the quadratic equation gives the correct spherical cap height:

$$h = \frac{-b_q - \sqrt{b_q^2 - 4a_q c_q}}{2a_q} \tag{D.6}$$

The spherical cap area is then simply:

$$S_{curve} = 2\pi R h(\theta_S) \tag{D.7}$$

Hence, each SCANDI zone's area is

$$A_{curve} = \frac{S_2 - S_1}{N} \tag{D.8}$$

where S_2 and S_1 denotes the spherical cap area covers by two consecutive SCANDI rings, and N is the number of zones in the SCANDI ring of interest.

The final results, for an ∼equal area projection scheme can be found in table 3.3.

D.2 SCANDI zone solid angles and flat projection

The solid angle of a cone or a spherical cap, Ω, is give by

$$\Omega = 2\pi(1 - cos\theta_s). \tag{D.9}$$

This is applicable to the SCANDI central zone where θ_s is the zone's zenith angle.

A more general equation for the solid angles of all the SCANDI zones can be written as

$$\Omega = \frac{2\pi(cos\theta_2 - cos\theta_1)}{N} \tag{D.10}$$

where θ_2 and θ_1 are the outer and inner zenith angles respectively, for a SCANDI ring, and N is the number of zones in the ring.

The projected elemental area of a zone on a flat sky dS_{flat} is given by

$$dS_{flat} = xd\theta dx \tag{D.11}$$

Therefore the total projected area S_{flat} becomes

$$S_{flat} = \int_{\theta_1}^{\theta_2} \int_{x_1}^{x_2} dS_{flat} \tag{D.12}$$

$$S_{flat} = \frac{\pi h_s^2 \left(tan^2\theta_2 - tan^2\theta_1\right)}{N} \tag{D.13}$$

where h_s is the height of the emission layer as before.

The flat area projection can also be found in table 3.3, where it is compared against the curved projected area.

Appendix E

Neutral temperature trend study

E.1 Green line emission, 557.7nm

Production rates

The green line emission arises from the transition of O(^{1}S) state to the O(^{1}D) state. The height profile of the volume emission rate can be calculated, and with the assumption of photochemical equilibrium, it can be expressed as

$$I_G = A_{557.7}[O(^1S)] = A_{557.7}\frac{\sum_{i'=1}^{n'} P_{i'}^{557.7}}{\sum_{j'=1}^{m'} L_{j'}^{557.7} + A(^1S)} \quad cm^{-3}s^{-1} \tag{E.1}$$

where

- I_G = the volume emission rate of the green-line emissions ($cm^{-3}s^{-1}$);
- $A_{557.7}$ = the Einstein transition probability for green-line emissions (s^{-1}). $A_{557.7}$ is used by Vlasov et al. [2005] and equals 1.215;
- $[O(^1S)]$ = the density of $O(^1S)$ at photoequilibrium (cm^{-3}), and the squared-bracket represents densities;
- $P_{i'}^{557.7}$ = the production rate of the processes involved in $O(^1S)$ productions ($cm^{-3}s^{-1}$);
- $L_{i'}^{557.7}$ = the loss rate of the processes involved in $O(^1S)$ collisional deactivation (s^{-1}); and finally
- $A(^1S)$ = the sum of probabilities for the emissions from $O(^1S)$ excited states in units of s^{-1}. $A(^1S)$ is applied in the green line model of Vlasov et al. [2005], and equals 1.291.

The green line emission has also been described in detail by many authors, although with slightly different conclusions as to which mechanisms should be included in the model e.g. Mullen et al. [1977], Witasse et al. [1999], Vlasov et al. [2005]. The study here follows the models in Witasse et al. [1999] and Vlasov et al. [2005], and only the $O(^1S)$ productions applicable to nightglow are discussed.

The first production source is the dissociative recombination of O_2^+:

$$O_2^+ + e_{th} \rightarrow O + O(^1S) \tag{E.2}$$

with the production rate:

$$P_{1'}^{557.7}(z) = \beta_{1'} k_{1'} [O_2^+](z) N_e(z) \tag{E.3}$$

where $\beta_{1'}$ is the yield and $k_{1'}$ is the rate coefficient. The $\beta_{1'} k_{1'}$ values chosen by Witasse et al. [1999] and Vlasov et al. [2005] are different. Witasse et al. [1999] considered the vibrational distribution of O_2^+, and gives

$$\beta_{1'} k_{1'} = \beta_{1'} * [0.9 * 8.1 \times 10^{-10} (\frac{1150}{T_e})^{-1.47} + 0.1 * 7.5 \times 10^{-8} (\frac{300}{T_e})^{0.34}], \tag{E.4}$$

but did not specify the $\beta_{1'}$ value in the paper. However, a different vibrational distribution is chosen by Vlasov et al. [2005]:

$$\beta_{1'} k_{1'} = 0.08 * 1.95 \times 10^{-7} (\frac{300}{T_e})^{0.7}. \tag{E.5}$$

There is also a small contribution from the following reaction:

$$N + O_2^+ \rightarrow NO^+ + O(^1S) \tag{E.6}$$

$$P_{2'}^{557.7} = \beta_{2'} k_{2'} [N][O_2^+](z). \tag{E.7}$$

However, this reaction is excluded in this study even though Witasse et al. [1999] consider this reaction as non-negligible, because the density of N is at least factor of 5 smaller than the major neutral species at 110km altitude.

The loss rates involved in equation E.1 are the collision deactivations of $O(^1S)$, which has a shorter radiative lifetime of 1s:

$$O(^1S) + X \rightarrow O + X \tag{E.8}$$

and only O_2 and O are considered, with the loss rates:

$$L_{1'}^{557.7} = k'_{O_2}[O_2](z) \tag{E.9}$$

$$L_{2'}^{557.7} = k'_{O}[O](z). \tag{E.10}$$

The final mechanism, which is expressed as a separate term from equation E.1, is the Barth two-step mechanism for mesospheric heights. Mullen et al. [1977] have made a detailed study of the mesospheric contribution to nightglow in the polar cap, and it considers both the Chapman three body mechanism, as well as Barth's, and states that their experimental result matches better with the Barth two-step mechanism:

$$O(^3P) + O(^3P) + M \rightarrow O_2^* + M \tag{E.11}$$

$$O_2^* + O(^3P) \rightarrow O_2 + O(^1S). \tag{E.12}$$

The production rate of the Barth mechanism is taken from Vlasov et al. [2005], and it is treated as a separate additional term to the volume emission rate from equation E.1:

$$I_G^{mes} = A_{557.7} \frac{k_1^{mes} k_2^{mes} [O(^3P)][M]}{(k_3^{mes}[O_2 + k_4^{mes}[O] + A_\Sigma)(k_5^{mes}[O_2] + A_{1S})} \tag{E.13}$$

where

- [M]=[(O_2)]+[(N_2)]
- A_Σ=0.02
- k_1^{mes}=$7.5 \times 10^{-34}(300/T_n)^2$
- k_2^{mes}=0.2*0.6$\times 10^{-12}$
- k_3^{mes}=3.0$\times 10^{-14}$
- k_4^{mes}=6.0$\times 10^{-12}$
- k_5^{mes}=4.3$\times 10^{-12} exp(-850/T_n)$

Since the SCANDI/FPIs data are night-time measurements only, the dayglow sources mentioned in Witasse et al. [1999] are excluded from the photo-emission model here, which includes day-time production sources of electron impact, collisional deactivation of $N_2(A^3\Sigma_u^+)$ or photodissociation of O_2 molecules.

The different production rates in the green line, unlike those of the red line, peak at different altitudes. The relevant nighttime contributions are discussed by Witasse et al. [1999]. There are approximately three peaks of different sizes for the green line nightglow. The first peak is a F-region peak from dissociative recombination, the second peak is a smaller peak due to dissociative recombination at 110km, with a small contribution from the N reaction, which is not included in the experiment. And the lowest peak is the mesospheric peak from the Barth mechanism.

This analysis models all three green emission peaks, which all should theoretically appear in the high latitude night time atmosphere. The overall green line emission profile has contributions from all three. It is generally accepted that the green line peak is broader and less well defined. Also please note this is a section included in the appendices for completeness to discuss the methodology of modelling the greenline emission profile. It is one of the emission lines observed by the FPI, but is not involved in the science experiments in this thesis.

Bibliography

V. J. Abreu. Atmospheric scattering effects on ground-based Fabry-Perot measurements of thermospheric winds - An inversion technique. *Planetary and Space Science*, 33: 1049–1055, September 1985. doi: 10.1016\0032-0633(85)90023-6.

V. J. Abreu, G. A. Schmitt, P. B. Hays, J. W. Meriwether, Jr., C. A. Tepley, and L. L. Cogger. Atmospheric scattering effects on ground-based measurements of thermospheric winds. *Planetary and Space Science*, 31:303–310, March 1983. doi: 10.1016\0032-0633(83)90080-6.

A. Aksnes, R. Eastes, S. Budzien, and K. Dymond. Dependence of neutral temperatures in the lower thermosphere on geomagnetic activity. *Journal of Geophysical Research (Space Physics)*, 112:6302–+, June 2007. doi: 10.1029\2006JA012214.

E. Alge, N. G. Adams, and D. Smith. Measurements of the dissociative recombination coefficients of O2(+), NO(+) and NH4(+) in the temperature range 200-600 K. *Journal of Physics B Atomic Molecular Physics*, 16:1433–1444, April 1983. doi: 10.1088/0022-3700/16/8/017.

A. L. Aruliah, A. D. Farmer, D. Rees, and U. Brändström. The seasonal behavior of high-latitude thermospheric winds and ion velocities observed over one solar cycle. *Journal of Geophysical Research*, 101:15701–15712, July 1996. doi: 10.1029\96JA00360.

A. L. Aruliah, I. C. F. Müller-Wodarg, and J. Schoendorf. Consequences of geomagnetic history on the high-latitude thermosphere and ionosphere: Averages. *Journal of Geophysical Research*, 104:28073–28088, 1999. doi: 10.1029\1999JA900334.

A. L. Aruliah, E. M. Griffin, I. McWhirter, A. D. Aylward, E. A. K. Ford, A. Charalambous, M. J. Kosch, Davis C. J., and Howells V. S. C. First tristatic study of meso-scale ion-neutral dynamics and energetics in the high-latitude upper atmosphere

using co-located FPIs and EISCAT radar. *Geophysical Research Letters*, 31:147–162, 2004.

A. L. Aruliah, E. M. Griffin, A. D. Aylward, E. A. K. Ford, M. J. Kosch, C. J. Davis, V. S. C. Howells, S. E. Pryse, H. R. Middleton, and J. Jussila. First direct evidence of meso-scale variability on ion-neutral dynamics using co-located tristatic FPIs and EISCAT radar in Northern Scandinavia. *Annales Geophysicae*, 23:147–162, January 2005.

A. L. Aruliah, E. M. Griffin, H.-C. I. Yiu, I. McWhirter, A. Charalambous, I. W. McCrea, and V. S. C. Howells. Mesoscale structure in ion-neutral coupling over Svalbard. *EGU General Assembly 2009, held 19-24 April, 2009 in Vienna, Austria, p.8782*, 11:8782–+, April 2009.

A. L. Aruliah, E. M. Griffin, H.-C. I. Yiu, I. McWhirter, and A. Charalambous. SCANDI - an all-sky Doppler imager for studies of thermospheric spatial structure. *Annales Geophysicae*, 28:549–567, February 2010.

A. L. Aruliah, A. Ronksley, H. C. Carlson, I. McWhirter, T. C. Spain, I. Meriwether, and I. McCrea. Large, sustained, vertical winds observed in the thermosphere during cusp heating/precipitation by FPIs and EISCAT radar on Svalbard. *Journal of Geophysical Research*, 2013.

S. M. I. Azeem, E. R. Talaat, G. G. Sivjee, and J.-H. Yee. Mesosphere and lower thermosphere temperature anomalies during the 2002 Antarctic stratospheric warming event. *Annales Geophysicae*, 28:267–276, January 2010.

S. Batten, D. Rees, D. Wade, and A. Steen. Observations of thermospheric neutral winds by the UCL Doppler imaging system at Kiruna in northern Scandinavia. *Journal of Atmospheric and Terrestrial Physics*, 50:861–888, November 1988.

P. R. Bevington. *Data reduction and error analysis for the physical sciences*. 1969.

D. Bilitza, D. Rawer, L. Bossy, and T. Guliaeva. International Reference Ionosphere: Past, present, and future. I - Electron density. II - Plasma temperatures, ion composition and ion drift. *Advances in Space Research*, 13:3–13, March 1993. doi: 10.1016\0273-1177(93)90240-C.

A. Brekke, J. R. Doupnik, and P. M. Banks. Incoherent Scatter Measurements of E Region Conductivities and Currents in the Auroral Zone. *Journal of Geophysical Research*, 79: 3773–3790, 1974. doi: 10.1029\JA079i025p03773.

K. A. Browning and R. Wexler. The Determination of Kinematic Properties of a Wind Field Using Doppler Radar. *Journal of Applied Meteorology*, 7:105–113, February 1968. doi: 10.1175\1520-0450(1968)007⟨0105:TDOKPO⟩2.0.CO;2.

J. K. Burchill, D. J. Knudsen, J. H. Clemmons, K. Oksavik, R. F. Pfaff, C. T. Steigies, A. W. Yau, and T. K. Yeoman. Thermal ion upflow in the cusp ionosphere and its dependence on soft electron energy flux. *Journal of Geophysical Research (Space Physics)*, 115:5206–+, May 2010. doi: 10.1029\2009JA015006.

R. G. Burnside, F. A. Herrero, J. W. Meriwether, Jr., and J. C. G. Walker. Optical observations of thermospheric dynamics at Arecibo. *Journal of Geophysical Research*, 86:5532–5540, July 1981. doi: 10.1029\JA086iA07p05532.

H. C. Carlson. Role of neutral atmospheric dynamics in cusp density and ionospheric patch formation. *Geophysical Research Letters*, 341:L13101, July 2007. doi: 10.1029/2007GL029316.

H. C. Carlson, J. Moen, K. Oksavik, C. P. Nielsen, I. W. McCrea, T. R. Pedersen, and P. Gallop. Direct observations of injection events of subauroral plasma into the polar cap. *Geophysical Research Letters*, 33:L05103, March 2006. doi: 10.1029/2005GL025230.

S. Chandra and A. K. Sinha. The diurnal heat budget of the thermosphere. *Planetary and Space Science*, 21:593–604, April 1973. doi: 10.1016\0032-0633(73)90072-X.

G. Chisham, M. Pinnock, A. S. Rodger, and J.-P. Villain. High-time resolution conjugate SuperDARN radar observations of the dayside convection response to changes in IMF B y. *Annales Geophysicae*, 18:191–201, February 2000. doi: 10.1007\s00585-000-0191-y.

J. H. Clemmons, J. H. Hecht, D. R. Salem, and D. J. Strickland. Thermospheric density in the Earth's magnetic cusp as observed by the Streak mission. *Geophysical Research Letters*, 35:24103–+, December 2008. doi: 10.1029\2008GL035972.

M. Conde. Deriving wavelength spectra from fringe images from a fixed-gap single-etalon Fabry-Perot spectrometer. *Applied Optics*, 41:2672–2678, May 2002.

M. Conde and R. W. Smith. Phase compensation of a separation scanned, all-sky imaging Fabry Perot spectrometer for auroral studies. *Applied Optics*, 36:5441–5450, August 1997. doi: 10.1364\AO.36.005441.

M. Conde and R. W. Smith. Spatial structure in the thermospheric horizontal wind above Poker Flat, Alaska, during solar minimum. *Journal of Geophysical Research*, 103:9449–9472, May 1998. doi: 10.1029\97JA03331.

C. S. Deehr, G. G. Sivjee, A. Egeland, K. Henriksen, P. E. Sandholt, R. Smith, P. Sweeney, C. Duncan, and J. Gilmer. Ground-based observations of F region aurora associated with the magnetospheric cusp. *Journal of Geophysical Research*, 85:2185–2192, May 1980. doi: 10.1029\JA085iA05p02185.

H. G. Demars and R. W. Schunk. Thermospheric response to ion heating in the dayside cusp. *Journal of Atmospheric and Solar-Terrestrial Physics*, 69:649–660, April 2007. doi: 10.1016\j.jastp.2006.11.002.

T. Dudok de Wit and S. Bruinsma. Determination of the most pertinent EUV proxy for use in thermosphere modeling. *Geophysical Research Letters*, 381:L19102, October 2011. doi: 10.1029/2011GL049028.

J. T. Emmert, J. L. Lean, and J. M. Picone. Record-low thermospheric density during the 2008 solar minimum. *Geophysical Research Letters*, 37:12102–+, June 2010. doi: 10.1029\2010GL043671.

L. Floyd, D. McMullin, and L. Herring. Intercomparison of SEM irradiances and solar indices. In A. Wilson, editor, *From Solar Min to Max: Half a Solar Cycle with SOHO*, volume 508 of *ESA Special Publication*, pages 197–200, June 2002.

J. C. Foster, J. R. Doupnik, and G. S. Stiles. Large scale patterns of auroral ionospheric convection observed with the Chatanika radar. *Journal of Geophysical Research*, 86: 11357–11371, December 1981. doi: 10.1029\JA086iA13p11357.

J. C. Foster, J. M. Holt, R. G. Musgrove, and D. S. Evans. Ionospheric convection associated with discrete levels of particle precipitation. *Geophysical Research Letters*, 13: 656–659, July 1986. doi: 10.1029\GL013i007p00656.

L. Goncharenko and S.-R. Zhang. Ionospheric signatures of sudden stratospheric warming: Ion temperature at middle latitude. *Geophysical Research Letters*, 35:21103–+, November 2008. doi: 10.1029\2008GL035684.

L. Goncharenko, J. L. Chau, P. Condor, A. Coster, and L. Benkevitch. Ionospheric effects of sudden stratospheric warming during moderate-to-high solar activity: Case study of January 2013. *Geophysical Research Letters*, 40:4982–4986, October 2013. doi: 10.1002/grl.50980.

E. M. Griffin. Comparison of high latitude thermospheric meridional neutral wind climatologies. *PhD Thesis, Department of Physics and Astonomy, University College London, University of London*, July 2000.

E. M. Griffin, A. L. Aruliah, I. McWhirter, H.-C. I. Yiu, A. Charalambous, and I. McCrea. Upper thermospheric neutral wind and temperature measurements from an extended spatial field. *Annales Geophysicae*, 26:2649–2655, September 2008. doi: 10.5194\angeo-26-2649-2008.

E. M. Griffin, A. L. Aruliah, I. McWhirter, H.-C. I. Yiu, and A. Charalambous. Upper thermospheric ion-neutral coupling from combined optical and radar experiments over Svalbard. *Annales Geophysicae*, 27:4293–4303, November 2009. doi: 10.5194\angeo-27-4293-2009.

B. Gustavsson, B. U. E. Brändström, Å. Steen, T. Sergienko, T. B. Leyser, M. T. Rietveld, T. Aso, and M. Ejiri. Nearly simultaneous images of HF-pump enhanced airglow at 6300 Å and 5577 Å. *Geophysical Research Letters*, 29(24):2220, December 2002. doi: 10.1029/2002GL015350.

J. K. Hargreaves. The solar-terrestrial environment. An introduction to geospace - the science of the terrestrial upper atmosphere, ionosphere and magnetosphere. *Camb. Atmos. Space Sci. Ser., Vol. 5,*, 5, 1992.

M. J. Harris. A new coupled terrestrial mesosphere-thermosphere general circulation model : Studies of dynamic, energetic, and photochemical coupling in the middle and upper atmosphere. *PhD Thesis, University of London, UK*, 2001.

D. H. Hathaway and M. Adams. The Butterfly Diagram. *Solar Physics, Marshall Space Flight Center, National Aeronautics and Space Administration (NASA)*, October 2010.

P. B. Hays, D. W. Rusch, R. G. Roble, and J. C. G. Walker. The O I (6300 Å) Airglow. *Reviews of Geophysics*, 16:225–232, 1978. doi: 10.1029\RG016i002p00225.

D. F. Heath and B. M. Schlesinger. The Mg 280-nm doublet as a monitor of changes in solar ultraviolet irradiance. *Journal of Geophysical Research*, 91:8672–8682, July 1986. doi: 10.1029\JD091iD08p08672.

A. E. Hedin. MSIS-86 thermospheric model. *Journal of Geophysical Research*, 92:4649–4662, May 1987. doi: 10.1029\JA092iA05p04649.

A. E. Hedin. Extension of the MSIS thermosphere model into the middle and lower atmosphere. *Journal of Geophysical Research*, 96:1159–1172, February 1991. doi: 10.1029\90JA02125.

R. A. Heelis, D. McEwen, and W. Guo. Ion and neutral motions observed in the winter polar upper atmosphere. *Journal of Geophysical Research (Space Physics)*, 107:1476–+, December 2002. doi: 10.1029\2002JA009359.

R. A. Heelis, W. R. Coley, A. G. Burrell, M. R. Hairston, G. D. Earle, M. D. Perdue, R. A. Power, L. L. Harmon, B. J. Holt, and C. R. Lippincott. Behavior of the O+$\ddot{+}$ transition height during the extreme solar minimum of 2008. *Geophysical Research Letters*, 36:0–+, July 2009. doi: 10.1029\2009GL038652.

J. M. Holt and S. R. Zhang. Long-term temperature trends in the ionosphere above Millstone Hill. *Geophysical Research Letters*, 35:5813–+, March 2008. doi: 10.1029\2007GL031148.

IDL. *IDL Reference Guide, IDL Version 6.3, April 2006 Edition, IDL Online Help (March 01, 2006)*. RSI, 2006.

M. C. Kelly. *The Earth's Ionosphere: Plasma Physics and Electrodynamics*. Academic Press, 1989.

T. L. Killeen and R. G. Roble. An analysis of the high-latitude thermospheric wind pattern calculated by a thermospheric general circulation model. I - Momentum forcing. *Journal of Geophysical Research*, 89:7509–7522, September 1984. doi: 10.1029\JA089iA09p07509.

T. L. Killeen, Y.-I. Won, R. J. Niciejewski, and A. G. Burns. Upper thermosphere winds and temperatures in the geomagnetic polar cap: Solar cycle, geomagnetic activity, and interplanetary magnetic field dependencies. *Journal of Geophysical Research*, 100:21327–21342, November 1995. doi: 10.1029\95JA01208.

M. J. Kosch, C. Anderson, R. A. Makarevich, B. A. Carter, R. A. D. Fiori, M. Conde, P. L. Dyson, and T. Davies. First E region observations of mesoscale neutral wind interaction with auroral arcs. *Journal of Geophysical Research (Space Physics)*, 115: 2303–+, February 2010. doi: 10.1029\2009JA014697.

M. J. Kosch, H.-C. I. Yiu, C. Anderson, T. Tsuda, Y. Ogawa, S. Nozawa, A. Arulliah, V. Howells, L. Baddeley, I. McCrea, and J. Wild. Meso-scale observations of Joule heating near an auroral arc and ion-neutral collision frequency in the polar cap E-region. *Journal of Geophysical Research (Space Physics)*, 2011.

J. (. Kumler and M. L. Bauer. Fish-eye lens designs and their relative performance. In R. E. Fischer, R. B. Johnson, W. J. Smith, & W. H. Swantner, editor, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 4093 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 360–369, October 2000.

J. Laštovička, R. A. Akmaev, G. Beig, J. Bremer, J. T. Emmert, C. Jacobi, M. J. Jarvis, G. Nedoluha, Y. I. Portnyagin, and T. Ulich. Emerging pattern of global change in the upper atmosphere and ionosphere. *Annales Geophysicae*, 26:1255–1268, May 2008.

J. L. Lean, T. N. Woods, F. G. Eparvier, R. R. Meier, D. J. Strickland, J. T. Correira, and J. S. Evans. Solar extreme ultraviolet irradiance: Present, past, and future. *Journal of Geophysical Research (Space Physics)*, 116:A01102, January 2011. doi: 10.1029\2010JA015901.

T. B. Leyser, B. Gustavsson, B. U. E. Brändström, E. Steen, F. Honary, M. T. Rietveld, T. Aso, and M. Ejiri. Enhanced Airglow by High Frequency Electromagnetic Pumping of the Ionosphere at Auroral Latitudes. *APS Meeting Abstracts*, pages 1037P–+, October 2000.

R. Link and L. L. Cogger. A reexamination of the O I 6300-A nightglow. *Journal of Geophysical Research*, 93:9883–9892, September 1988. doi: 10.1029\JA093iA09p09883.

M. Lockwood and H. C. Carlson, Jr. Production of polar cap electron density patches by transient magnetopause reconnection. *Geophysical Research Letters*, 19:1731–1734, September 1992. doi: 10.1029/92GL01993.

M Lockwood, I.W McCrea, G.H Millward, R.J Moffett, and H Rishbeth. Eiscat observations of ion composition and temperature anisotropy in the high-latitude f-region. *Journal of Atmospheric and Terrestrial Physics*, 55(6):895 – 906, 1993. ISSN 0021-9169. doi: DOI:10.1016\0021-9169(93)90029-X.

G. Lu, A. D. Richmond, B. A. Emery, and R. G. Roble. Magnetosphere-ionosphere-thermosphere coupling: Effect of neutral winds on energy transfer and field-aligned current. *Journal of Geophysical Research*, 100:19643–19660, October 1995. doi: 10.1029\95JA00766.

H. Lühr, M. Rother, W. Köhler, P. Ritter, and L. Grunwaldt. Thermospheric up-welling in the cusp region: Evidence from CHAMP observations. *Geophysical Research Letters*, 31:6805–+, March 2004. doi: 10.1029\2003GL019314.

H. G. Mayr and H. Volland. Theoretical model for the latitude dependence of the thermospheric annual and semiannual variations. *J. Geophys. Res.*, 77:6774–6790, 1972. doi: 10.1029\JA077i034p06774.

H. G. Mayr, I. Harris, F. A. Herrero, and F. Varosi. Winds and composition changes in the thermosphere using the transfer function model. *Journal of Atmospheric and Solar-Terrestrial Physics*, 59:691–709, April 1997. doi: 10.1016\S1364-6826(96)00100-9.

F. G. McCormac, T. L. Killeen, B. Nardi, and R. W. Smith. How close are ground-based Fabry-Perot Thermospheric wind and temperature measurements to exospheric values? A simulation study. *Planet. Space Sci.*, 10:1255–1265, May 1987.

F. G. McCormac, T. L. Killeen, A. G. Burns, J. W. Meriwether, Jr., R. G. Roble, L. E. Wharton, and N. W. Spencer. Polar cap diurnal temperature variations - Observations and modeling. *Journal of Geophysical Research*, 93:7466–7477, July 1988. doi: 10.1029\JA093iA07p07466.

I. W. McCrea, M. Lester, T. R. Robinson, J.-P. St.-Maurice, N. M. Wade, and T. B. Jones. Derivation of the ion temperature partition coefficient beta-parallel from the study of

ion frictional heating events. *Journal of Geophysical Research*, 98:15701–+, September 1993. doi: 10.1029\92JA02776.

I. McWhirter. Imaging Photon Detectors and their use with Single and Multiple Fabry-Perot Etalon Systems for Atmospheric Wind Measurements. *PhD Thesis, Department of Physics and Astonomy, University College London, University of London*, 1993.

I. McWhirter, E. M. Griffin, A. L. Aruliah, E. A. K. Ford, A. Charalambous, and M. J. Kosch. Optical Instrumentation of the Atmospheric Physics Laboratory University College London. *Proceedings of the 17th ESA Symposium on European Rocket and Balloon Programs and Related Research, Sandefjord, Norway, 30 May - 2 June 2005*, pages 315–319, August 2005.

J. Moen, K. Oksavik, and H. C. Carlson. On the relationship between ion upflow events and cusp auroral transients. *Geophysical Research Letters*, 31:L11808, June 2004. doi: 10.1029/2004GL020129.

E. G. Mullen, S. M. Silverman, and D. F. Korff. Nightglow (557.7 nm of OI) in the central polar cap. *Planetary Space Science*, 25:23–38, January 1977. doi: 10.1016\0032-0633(77)90114-3.

NASA GSFC and D. Bilitza. International Reference Ionosphere. *Heliospheric Physics Lab*, November 2010.

P. T. Newell and C.-I. Meng. The cusp and the cleft boundary layer - Low-altitude identification and statistical local time variation. *Journal of Geophysical Research*, 93: 14549–14556, December 1988. doi: 10.1029\JA093iA12p14549.

P. T. Newell and C.-I. Meng. Dipole tilt angle effects on the latitude of the cusp and cleft low-latitude boundary layer. *Journal of Geophysical Research*, 94:6949–6953, June 1989. doi: 10.1029\JA094iA06p06949.

P. T. Newell, T. Sotirelis, J. M. Ruohoniemi, J. F. Carbary, K. Liou, J. P. Skura, C.-I. Meng, C. Deehr, D. Wilkinson, and F. J. Rich. OVATION: Oval variation, assessment, tracking, intensity, and online nowcasting. *Annales Geophysicae*, 20:1039–1047, July 2002. doi: 10.5194\angeo-20-1039-2002.

P. T. Newell, T. Sotirelis, K. Liou, A. R. Lee, S. Wing, J. Green, and R. Redmon. Predictive ability of four auroral precipitation models as evaluated using Polar UVI global images. *Space Weather*, 8:S12004, December 2010. doi: 10.1029\2010SW000604.

National Oceanic Atmospheric Administration NOAA. Geomagnetic kp and ap Indices. *www.ngdc.noaa.gov*, 2011.

Y. Ogawa, S. C. Buchert, R. Fujii, S. Nozawa, and A. P. van Eyken. Characteristics of ion upflow and downflow observed with the European Incoherent Scatter Svalbard radar. *Journal of Geophysical Research (Space Physics)*, 114:5305–+, May 2009. doi: 10.1029\2008JA013817.

JHU APL OVATION. (Auroral) Oval Variation, Assessment, racking, Intensity, and Online Nowcasting. *www.jhuapl.edu\Aurora\ovation*, 2011.

J. M. Picone, A. E. Hedin, D. P. Drob, and A. C. Aikin. NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues. *Journal of Geophysical Research (Space Physics)*, 107:1468, December 2002. doi: 10.1029\2002JA009430.

W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, and H. Gould. Numerical Recipes, The Art of Scientific Computing. *American Journal of Physics*, 55:90–91, January 1987. doi: 10.1119/1.14981.

K. Rawer, D. Bilitza, and S. Ramakrishnan. Goals and status of the International Reference Ionosphere. *Reviews of Geophysics and Space Physics*, 16:177–181, May 1978.

M. H. Rees and R. A. Jones. Time dependent studies of the aurora-II. Spectroscopic morphology. *Planetary and Space Science*, 21:1213–1235, July 1973. doi: 10.1016\0032-0633(73)90207-9.

M. H. Rees and D. Luckey. Auroral electron energy derived from ratio of spectroscopic emissions. I - Model computations. *Journal of Geophysical Research*, 79:5181–5186, December 1974. doi: 10.1029\JA079i034p05181.

M. H. Rees, K. Stamnes, B. A. Emery, and R. G. Roble. Neutral and ion gas heating by auroral electron precipitation. *Journal of Geophysical Research*, 88:6289–6300, August 1983. doi: 10.1029\JA088iA08p06289.

C. Reigber, H. Lühr, and P. Schwintzer. CHAMP mission status. *Advances in Space Research*, 30:129–134, July 2002. doi: 10.1016\S0273-1177(02)00276-4.

A. D. Richmond, C. Lathuillère, and S. Vennerstroem. Winds in the high-latitude lower thermosphere: Dependence on the interplanetary magnetic field. *Journal of Geophysical Research (Space Physics)*, 108:1066, February 2003. doi: 10.1029\2002JA009493.

H. Rishbeth, K. J. F. Sedgemore-Schulthess, and T. Ulich. Semiannual and annual variations in the height of the ionospheric F2-peak. *Annales Geophysicae*, 18:285–299, March 2000. doi: 10.1007\s005850050889.

R. G. Roble and E. C. Ridley. An auroral model for the NCAR thermospheric general circulation model (TGCM). *Annales Geophysicae*, 5:369–382, December 1987.

R. G. Roble, E. C. Ridley, and R. E. Dickinson. On the global mean structure of the thermosphere. *Journal of Geophysical Research*, 92:8745–8758, August 1987. doi: 10.1029\JA092iA08p08745.

A. S. Rodger, S. B. Mende, T. J. Rosenberg, and K. B. Baker. Simultaneous optical and HF radar observations of the ionospheric cusp. *Geophysical Research Letters*, 22:2045–2048, 1995. doi: 10.1029/95GL01797.

P.-E. Sandholt, K. Henriksen, C. S. Deehr, G. G. Sivjee, G. J. Romick, and A. Egeland. Dayside cusp auroral morphology related to nightside magnetic activity. *Journal of Geophysical Research*, 85:4132–4138, August 1980. doi: 10.1029\JA085iA08p04132.

P. E. Sandholt, C. S. Deehr, A. Egeland, B. Lybekk, and R. Viereck. Signatures in the dayside aurora of plasma transfer from the magnetosheath. *Journal of Geophysical Research*, 91:10063–10079, September 1986. doi: 10.1029\JA091iA09p10063.

K. Schlegel, H. Lühr, J.-P. S. -Maurice, G. Crowley, and C. Hackert. Thermospheric density structures over the polar regions observed with CHAMP. *Annales Geophysicae*, 23:1659–1672, July 2005. doi: 10.5194\angeo-23-1659-2005.

G. Schubert, R. L. Walterscheid, M. P. Hickey, and C. A. Tepley. Observations and interpretation of gravity wave induced fluctuations in the O I (557.7 nm) airglow. *Journal of Geophysical Research*, 104:14915–14924, July 1999. doi: 10.1029\1999JA900096.

R. W. Schunk. Transport equations for aeronomy. *Planetary and Space Science*, 23:437–485, March 1975. doi: 10.1016\0032-0633(75)90118-X.

R. W. Schunk and A. F. Nagy. *Ionospheres; Physics, Plasma Physics, and Chemistry*. Cambridge University Press, 2000a.

R. W. Schunk and A. F. Nagy. *Ionospheres; Physics, Plasma Physics, and Chemistry*, pages 16–21. Cambridge University Press, 2000b.

R. W. Schunk and A. F. Nagy. *Ionospheres; Physics, Plasma Physics, and Chemistry*, pages 404–406. Cambridge University Press, 2000c.

R. W. Schunk and A. F. Nagy. *Ionospheres; Physics, Plasma Physics, and Chemistry*, pages 289–289. Cambridge University Press, 2000d.

R. W. Schunk and A. F. Nagy. *Ionospheres; Physics, Plasma Physics, and Chemistry*, pages 129–136. Cambridge University Press, 2000e.

C. H. Sheehan and J.-P. St.-Maurice. Dissociative recombination of N_2^+, O_2^+, and NO^+: Rate coefficients for ground state and vibrationally excited ions. *Journal of Geophysical Research (Space Physics)*, 109:A03302, March 2004. doi: 10.1029/2003JA010132.

F. Sigernes. The Meridian Scanning Photometer. *http:\\fred.unis.no\6dec2002\msp.htm*, 2002.

G. G. Sivjee, G. J. Romick, and C. S. Deehr. Optical signatures of some magnetospheric processes on the dayside. *Geophysical Research Letters*, 9:676–679, June 1982. doi: 10.1029\GL009i006p00676.

M. F. Smith and M. Lockwood. Earth's magnetospheric cusps. *Reviews of Geophysics*, 34:233–260, 1996. doi: 10.1029\96RG00893.

R. W. Smith and G. Hernandez. Upper thermospheric temperatures at South Pole. *Advances in Space Research*, 16:31–, 1995. doi: 10.1016\0273-1177(95)00170-J.

R. W. Smith, K. Henriksen, C. S. Deehr, G. G. Sivjee, D. Rees, and F. G. McCormac. Thermospheric winds in the cusp - Dependence of the latitude of the cusp. *Planetary and Space Science*, 33:305–313, March 1985. doi: 10.1016\0032-0633(85)90062-5.

R. W. Smith, G. Hernandez, R. G. Roble, P. L. Dyson, M. Conde, R. Crickmore, and M. Jarvis. Observation and simulations of winds and temperatures in the Antarctic thermosphere for August 2-10, 1992. *Journal of Geophysical Research*, 103:9473–9480, May 1998. doi: 10.1029\97JA03336.

S. C. Solomon, P. B. Hays, and V. J. Abreu. The auroral 6300 A emission - Observations and modeling. *Journal of Geophysical Research*, 93:9867–9882, September 1988. doi: 10.1029\JA093iA09p09867.

K. F. Tapping. Recent solar radio astronomy at centimeter wavelengths - The temporal variability of the 10.7-cm flux. *Journal of Geophysical Research*, 92:829–838, January 1987. doi: 10.1029\JD092iD01p00829.

M. Thoby. `Fisheyelensescompared:Sigma8mmf\4vs.Nikkor10\5mmf\2.8.` *`http:\\michel.thoby.free.fr\Blur_Panorama\Nikkor10-5mm_or_Sigma8mm\Sigma_or_Nikkor\Comparison_Short_Version_Eng.html`*, June 2006.

A. P. Thorne. Spectrophysics 2^{nd} edition. *Chapman and Hall, London*, 1988.

W. K. Tobiska, S. D. Bouwer, and B. R. Bowman. The development of new solar indices for use in thermospheric density modeling. *Journal of Atmospheric and Solar-Terrestrial Physics*, 70:803–819, March 2008. doi: 10.1016/j.jastp.2007.11.001.

T. Ulich and E. Turunen. Evidence for long-term cooling of the upper atmosphere in ionosonde data. *Geophysical Research Letters*, 24:1103–1106, May 1997. doi: 10.1029\97GL50896.

R. A. Viereck, L. E. Floyd, P. C. Crane, T. N. Woods, B. G. Knapp, G. Rottman, M. Weber, L. C. Puga, and M. T. DeLand. A composite Mg II index spanning from 1978 to 2003. *Space Weather*, 2:S10005, October 2004. doi: 10.1029\2004SW000084.

N. V. Vlasov, M. J. Nicolls, M. C. Kelley, M. S. Smith, N. Aponte, and S. A. Gonzalez. Modeling of airglow and ionospheric parameters at Arecibo during quiet and disturbed periods in October 2002. *Journal of Geophysical Research*, 110, July 2005.

I. K. Walker, J. Moen, L. Kersley, and D. A. Lorentzen. On the possible role of cusp/cleft precipitation in the formation of polar-cap patches. *Annales Geophysicae*, 17:1298–1305, October 1999. doi: 10.1007/s00585-999-1298-4.

D. R. Weimer. Improved ionospheric electrodynamic models and application to calculating Joule heating rates. *Journal of Geophysical Research (Space Physics)*, 110:5306–+, May 2005. doi: 10.1029\2004JA010884.

O. Witasse, J. Lilensten, C. Lathuillère, and P.-L. Blelly. Modeling the OI 630.0 and 557.7 nm thermospheric dayglow during EISCAT-WINDII coordinated measurements. *Journal of Geophysical Research*, 104:24639–24656, November 1999. doi: 10.1029\1999JA900260.

X. Zhu, E. R. Talaat, J. B. H. Baker, and J.-H. Yee. A self-consistent derivation of ion drag and Joule heating for atmospheric dynamics in the thermosphere. *Annales Geophysicae*, 23:3313–3322, November 2005.